

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 01:23:55 ; Search time 44 Seconds
(without alignments)
2333.316 Million cell updates/sec

Title: US-09-834-271A-27

Perfect score: 185
Sequence: 1 ggccttaaggccctgcaatc.....tgtataagaataaggagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	3	US-09-384-305-22
2	185	100.0	185	3	US-09-258-377-27
3	183.4	99.1	185	3	US-09-384-305-21
4	183.4	99.1	185	3	US-09-258-377-26
5	180.2	97.4	185	3	US-09-384-305-20
6	180.2	97.4	185	3	US-09-258-377-25
7	158	85.4	10216	2	US-08-875-154-1
8	157.2	85.0	2604	3	US-09-264-097-3
9	157.2	85.0	2604	3	US-09-537-168-5
10	154	83.2	162	1	US-08-434-255-24
11	154	83.2	162	1	US-08-459-967-24
12	154	83.2	162	1	US-08-460-327-24
13	154	83.2	162	1	US-08-459-871-24
14	154	83.2	2084	1	US-08-720-859-3
15	154	83.2	2084	1	US-08-459-610-3
16	154	83.2	2084	2	US-08-343-804-3
17	154	83.2	2084	2	US-08-687-399-3
18	154	83.2	2084	2	US-08-600-908A-3
19	154	83.2	2084	3	US-08-683-838A-3
20	154	83.2	2084	3	US-09-182-859-3
21	154	83.2	2084	3	US-09-170-670-13
22	154	83.2	2084	3	US-09-193-088-31
23	154	83.2	2084	3	US-09-183-412-58
24	154	83.2	2084	4	US-09-290-734-13
25	154	83.2	2084	4	US-09-672-459-3
26	154	83.2	2084	4	US-09-636-252A-3
27	154	83.2	2084	4	US-09-545-586-13

28 154 83.2 2084 4 US-10-186-042-3 Sequence 3, Appli
29 154 83.2 2084 4 US-09-769-864-58 Sequence 58, Appli
30 44 23.8 44 3 US-09-384-305-15 Sequence 15, Appli
31 44 23.8 44 3 US-09-258-377-13 Sequence 19, Appli
32 39 21.1 42 2 US-08-873-479-37 Sequence 37, Appli
33 39 21.1 42 2 US-09-031-442A-9 Sequence 9, Appli
34 39 21.1 42 2 US-08-972-661A-25 Sequence 25, Appli
35 39 21.1 42 3 US-09-258-377-9 Sequence 9, Appli
36 38.4 20.8 43 2 US-08-873-479-38 Sequence 38, Appli
37 38.4 20.8 43 2 US-09-031-442A-10 Sequence 10, Appli
38 38.4 20.8 43 2 US-08-972-661A-26 Sequence 26, Appli
39 38.4 20.8 43 3 US-09-258-377-10 Sequence 10, Appli
40 34 18.4 64081 4 US-09-790-988-1 Sequence 1, Appli
41 33.2 17.9 1511 1 US-07-991-867B-8 Sequence 8, Appli
42 33.2 17.9 1511 1 US-08-107-755A-8 Sequence 8, Appli
43 33.2 17.9 1511 2 US-08-544-332-8 Sequence 8, Appli
44 33.2 17.9 1511 4 US-09-370-861A-8 Sequence 8, Appli
45 32.6 17.6 8700 2 US-08-392-625-16 Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-09-384-305-22
Sequence 22, Application US/09384305
Patent No. 6184028

GENERAL INFORMATION:
APPLICANT: Michael D. Thomas
TITLE OF INVENTION: Polypeptides Having Pectin
TITLE OF INVENTION: Acetylase Activity And Nucleic Acids Encoding Same
FILE REFERENCE: 5952,000-US
CURRENT APPLICATION NUMBER: US/09/384,305
CURRENT FILING DATE: 1999-08-26
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 22
LENGTH: 185
TYPE: DNA
ORGANISM: Bacillus subtilis
US-09-384-305-22

Query Match 100.0%; Score 185; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.3e-43;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCCTTAAGGGCTGCAATCGATTGTTTCAGAAAAGAGAGACCAATAAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTCAGAAAAGAGAGACCAATAAAATACCTTGT 60
Qy 61 CTGTCATCAGACAGGGGTATTTTATGCTGCTCCAGACTGCTCGCTGTGTAAAAAATAGGA 120
Db 61 CTGTCATCAGACAGGGGTATTTTATGCTGCTCCAGACTGCTCGCTGTGTAAAAAATAGGA 120
Qy 121 ATAAAGGGGGTTCAGATTATTTTACTGATATGATATATATATTTTGTATAGAAATGG 180
Db 121 ATAAAGGGGGTTCAGATTATTTTACTGATATGATATATATATTTTGTATAGAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 2
US-09-258-377-27
Sequence 27, Application US/09258377
Patent No. 6255075
GENERAL INFORMATION:
APPLICANT: Widner, William
APPLICANT: Sloma, Alan
APPLICANT: Thomas, Michael D.
TITLE OF INVENTION: Methods For Producing A polypeptide In a

```
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-27

Query Match      100.0%; Score 185; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.3e-43;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60

QY 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120
DB 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTATAAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTATAAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
US-09-384-305-21
; Sequence 21, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-21

Query Match      99.1%; Score 183.4; DB 3; Length 185;
Best Local Similarity 99.5%; Pred. No. 1.5e-42;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60

QY 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120
DB 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTATAAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTATAAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
US-09-258-377-26
; Sequence 26, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widheer, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-26

Query Match      99.1%; Score 183.4; DB 3; Length 185;
Best Local Similarity 99.5%; Pred. No. 1.5e-42;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60

QY 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120
DB 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTATAAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTATAAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
US-09-384-305-20
; Sequence 20, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-20

Query Match      97.4%; Score 180.2; DB 3; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.2e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
```

```
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
Qy 61 CTGTCTATCAGACAGGGTATTTTATGCTGCCAGACTGTCGGTGTGTAATAAATAGGA 120
Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGCCAGACTGTCGGTGTGTAATAAATAGGA 120
Qy 121 ATAAAGGGGGTGTGACATATTTTACTGATATGTAATAATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTGTGATATTTTACTGATATGTAATAATTTGTAAGAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 6
US-09-258-377-25
; Sequence 25, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Thomas, Alan
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; EARLIER FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-25

Query Match 97.4%; Score 180.2; DB 3; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.2e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGT 60
Qy 61 CTGTCTATCAGACAGGGTATTTTATGCTGCCAGACTGTCGGTGTGTAATAAATAGGA 120
Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGCCAGACTGTCGGTGTGTAATAAATAGGA 120
Qy 121 ATAAAGGGGGTGTGACATATTTTACTGATATGTAATAATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTGTGATATTTTACTGATATGTAATAATTTGTAAGAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 7
US-08-875-154-1
; Sequence 1, Application US/08875154
; Patent No. 5882888
; GENERAL INFORMATION:
; APPLICANT: Jorgensen, Streen Troels
; TITLE OF INVENTION: DNA Integration By Transporation
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5882888 No. 5882888disk of No. 5882888th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
```

```
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,154
FILING DATE: 17-JUL-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4381.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10216 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "PMOL553"
US-08-875-154-1

Query Match 85.4%; Score 158; DB 2; Length 10216;
Best Local Similarity 97.0%; Pred. No. 4.2e-35;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTATCAGACA 73
Db 2237 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTATCAGACA 2236
Qy 74 GGGTATTTTATGCTGCCAGACTGTCGGTGTGTAATAAATAGGAATAGGGGGT 133
Db 2297 GGGTATTTTATGCTGCCAGACTGTCGGTGTGTAATAAATAGGAATAGGGGGT 133
Qy 134 GACATTTTACTGATATGTAATAATTTGTAAGAAATG 179
Db 2357 GTTATTTTACTGATATGTAATAATTTGTAAGAAATG 2402

RESULT 8
US-09-264-097-3
; Sequence 3, Application US/09264097
; Patent No. 6287826
; GENERAL INFORMATION:
; APPLICANT: No. 6287826man, Barrie Edmund
; APPLICANT: Hendriksen, Hanne Vang
; TITLE OF INVENTION: Enzymatic Preparation of Glucose Syrup
; TITLE OF INVENTION: From Starch
; FILE REFERENCE: 5278.200-US
; CURRENT APPLICATION NUMBER: US/09/264,097
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: PA 0321/98
; EARLIER FILING DATE: 1998-03-09
; EARLIER APPLICATION NUMBER: 60/079,209
; EARLIER FILING DATE: 1998-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-09-264-097-3

Query Match 85.0%; Score 157.2; DB 3; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.2e-35;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTATCAGACAGGT 77
```

593 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGCTGTCATCAGACGGGT 652
78 ATTTTATCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAAGAGGGGGTGTGACA 137
653 ATTTTATCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAAGAGGGGGTGTGTTA 712
138 TTATTTACTGATATGATATATATATATTTGTATAGAAATG 179
713 TTATTTACTGATATGATATATATATTTGTATAGAAATG 754

RESULT 9

US-09-537-168-5
; Sequence 5, Application US/09537168
; Patent No. 6410295
; GENERAL INFORMATION:
; APPLICANT: Andersen, Carsten
; APPLICANT: Jorgensen, Christel T.
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjaerulf, Soren
; TITLE OF INVENTION: Alpha-Amylase Variants
; FILE REFERENCE: 5986.200-US
; CURRENT APPLICATION NUMBER: US/09/537,168
; EARLIER FILING DATE: 2000-03-29
; EARLIER APPLICATION NUMBER: PA 1999 00437
; EARLIER FILING DATE: 1999-03-30
; EARLIER APPLICATION NUMBER: 60/127,427
; EARLIER FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-09-537-168-5

Query Match 85.0%; Score 157.2; DB 4; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.2e-35;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGCTGTCATCAGACGGGT 77
DB 593 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGCTGTCATCAGACGGGT 652
QY 78 ATTTTATCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAAGAGGGGGTGTGACA 137
DB 653 ATTTTATCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAAGAGGGGGTGTGTTA 712
QY 138 TTATTTACTGATATGATATATATATTTGTATAGAAATG 179
DB 713 TTATTTACTGATATGATATATATTTGTATAGAAATG 754

RESULT 10

US-08-434-255-24
; Sequence 24, Application US/08434255
; Patent No. 5621089
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 5621089o No. 5621089disk of No. 5621089th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,255
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-434-255-24

Query Match 83.2%; Score 154; DB 1; Length 162;
Best Local Similarity 96.9%; Pred. No. 2.3e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGCTGTCATCAGACGGGT 77
DB 1 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGCTGTCATCAGACGGGT 60
QY 78 ATTTTATCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAAGAGGGGGTGTGACA 137
DB 61 ATTTTATCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAAGAGGGGGTGTGTTA 120
QY 138 TTATTTACTGATATGATATATATATTTGTATAGAAATG 179
DB 121 TTATTTACTGATATGATATATATTTGTATAGAAATG 162

RESULT 11

US-08-459-967-24
; Sequence 24, Application US/08459967
; Patent No. 5622841
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 5622841o No. 5622841disk of No. 5622841th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,967
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.

```

; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764,400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-459-967-24

Query Match      83.2%; Score 154; DB 1; Length 162;
Best Local Similarity 96.9%; Fred. No. 2.3e-3; Indels 0; Gaps 0
Matches 157; Conservative 0; Mismatches 5;

QY 18 ATCGATTGTTTCAGAAAAGAGACACCATAAAAATACCTTGTCTGTCATCAGACAGGT 77
Db 1 ATCGATTGTTTCAGAAAAGAGACACCATAAAAATACCTTGTCTGTCATCAGACAGGT 60

QY 78 ATTTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGAATAAGGGGGGTTGACA 137
Db 61 ATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGAATAAGGGGGGTTGTTA 120

QY 138 TTATTTTACTGATATGTAATAATAATTTGTAAGAAAATG 179
Db 121 TTATTTTACTGATATGTAATAATAATTTGTAAGAAAATG 162

RESULT 12
US-08-460-327-24
; Sequence 24, Application US/08460327
; Patent No. 5622850
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5622850o No. 5622850disk of No. 5622850th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,327
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764,400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

Db 61 ATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAGGAGGTTGTTA 120
QY 138 TTATTTTACTGATATGATATAATAATTTGTATAGAAAATG 179
Db 121 TTATTTTACTGATATGATATAATAATTTGTATAGAAAATG 162

RESULT 14

US-08-720-899-3
; Sequence 3, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5753460 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 250..1794
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 250..342
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 343..1791
US-08-720-899-3

Query Match 83.2%; Score 154; DB 1; Length 2084;
Best Local Similarity 96.9%; Pred. No. 3.9e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 18 ATCGATTGTTTGAGAAAAGAGAGACCATATAAATACCTTGTCTGTCATCAGACAGGTT 77
Db 73 ATCGATTGTTTGAGAAAAGAGAGACCATATAAATACCTTGTCTGTCATCAGACAGGTT 132
QY 78 ATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAGGAGTTGACA 137

Db 133 ATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAGGAGTTGTTA 192
QY 138 TTATTTTACTGATATGATATAATAATTTGTATAGAAAATG 179
Db 193 TTATTTTACTGATATGATATAATAATTTGTATAGAAAATG 234

RESULT 15

US-08-459-610-3
; Sequence 3, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5801043 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 250..1794
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 250..342
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 343..1791
US-08-459-610-3

Query Match 83.2%; Score 154; DB 1; Length 2084;
Best Local Similarity 96.9%; Pred. No. 3.9e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 18 ATCGATTGTTTGAGAAAAGAGAGACCATATAAATACCTTGTCTGTCATCAGACAGGTT 77
Db 73 ATCGATTGTTTGAGAAAAGAGAGACCATATAAATACCTTGTCTGTCATCAGACAGGTT 132
QY 78 ATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAGGAGTTGACA 137
Db 133 ATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAGGAGTTGTTA 192

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 01:21:15 ; Search time 1837 Seconds
(without alignments)
3007.351 Million cell updates/sec

Title: US-09-834-271A-27

Perfect score: 185

Sequence: 1 ggccctaaggcctgaatc.....tgtataagaaaatggagctc 185

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_lnv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssl:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43.4	23.5	427	14	CA819917
2	39.8	21.5	1201	13	EX356191
3	39.4	21.3	1201	9	AL554700
4	38.8	21.0	735	29	CE808066

C	5	38.6	20.9	1201	13	EX446359
C	6	38.6	20.9	1304	28	CC317939
C	7	37.6	20.3	392	10	AW292419
C	8	37.6	20.3	461	9	AW001963
C	9	37.6	20.3	462	12	BQ000230
C	10	37.6	20.3	467	9	AI989386
C	11	37.6	20.3	475	9	AI697642
C	12	37.6	20.3	564	9	AI479617
C	13	37.6	20.3	699	28	BZ072552
C	14	37.6	20.3	700	9	AU253100
C	15	37.4	20.2	894	9	AL543011
C	16	37	20.0	1010	10	BE614304
C	17	37	20.0	1038	29	CNS01TL7
C	18	36.8	19.9	700	28	BH922535
C	19	36.6	19.8	403	9	AI499797
C	20	36.6	19.8	404	9	AI570619
C	21	36.6	19.8	708	29	EX240763
C	22	36.6	19.8	822	29	CNS03LLL
C	23	36.4	19.7	792	29	CNS02D0S
C	24	36.4	19.7	1068	29	CNS00ETV
C	25	36.4	19.7	1200	13	EX426107
C	26	36.4	19.7	1201	13	EX442065
C	27	36.2	19.6	928	29	CNS00DKY
C	28	36	19.5	565	12	BM153720
C	29	36	19.5	650	28	BZ361891
C	30	36	19.5	695	28	BZ374571
C	31	36	19.5	938	29	CG288637
C	32	36	19.5	997	29	CG288649
C	33	36	19.5	1201	13	EX385059
C	34	35.8	19.4	498	13	EX569556
C	35	35.6	19.2	498	13	EX564726
C	36	35.6	19.2	504	13	EX568933
C	37	35.6	19.2	512	13	EX568933
C	38	35.6	19.2	534	12	BI371578
C	39	35.6	19.2	639	28	AZ362757
C	40	35.6	19.2	868	28	AZ535134
C	41	35.6	19.2	887	28	AZ546631
C	42	35.6	19.2	915	29	CG184658
C	43	35.6	19.2	995	29	CG184655
C	44	35.6	19.2	1201	13	EX419991
C	45	35.4	19.1	315	12	BG627035

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
CA819917	CA819917	sau82a05.y1 Gm-cl048 Glycine max	CA819917	CA819917	EST.	Glycine max (soybean)	Glycine max.
		Gm-cl048-2314 5', mRNA sequence.					Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
							1 (bases 1 to 427)

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Fape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
1. .427
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl048-2314"
/cissue type="whole seedlings of greenhouse grown plants"
/dev stage="1 week old"
/lab_host="DH10B"
/clone_lib="Gm-cl048"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The Clark NIL was constructed and seed was provided
by Dr. J. Specht, University of Nebraska (Shoemaker and
Specht, 1995). The cDNA library was constructed from mRNA
isolated from whole seedlings of 1 week old greenhouse
grown plants. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site and a 3' anchor. EcoRI adapters were
ligated to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (GibcoBRL). The library was constructed
in cooperation with Dr. Paul Kelm's laboratory at Northern
Arizona University."

FEATURES

```

1. 427
source
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GOYBEAN CLONE ID: Gm-c1048-2314"
/tissue_type="whole seedlings of greenhouse grown plants"
/dev_stage="1 week old"
/lab_host="DH10B"
/clone_lib="Gm-c1048"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The Clark NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from whole seedlings of 1 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). The library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."
ORIGIN
Query Match 23.5%; Score 43.4; DB 14; Length 427;
Best Local Similarity 53.9%; Pred. No. 1.2;
Matches 89; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 19 TCgATGCTTTGAGAAAAGAGACCATAAATACCTGTCTGTCATCAGACAGGTA 78
Db 164 TTGCCTTTTTTTAAAGGGAAAGGGGAAAACCCCTTTTTTTTAACTC 223
QY 79 TTTTATGCTGTCACAGACTGTCGCTGTGTAAAAAATAGGAATAAGGGGGTTGACAT 138
Db 224 TTTTCCAATCCCCCTTTTTTCCAAGAGGAAAAAAGAAAAAGCCCCCTTTAT 283
QY 139 TATTTTACTGATATGTAATAATAATTTGTATAGAAATGGAGC 183
Db 284 TTTTATTTAAATTTTTTATTTCAATTTTATAGAAATGGACC 328
RESULT 2
BX356191 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356191 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CSDD1009YC24 5-PRIME, mRNA sequence.
ACCESSION BX356191
VERSION BX356191.1 GI:30365992
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

```

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1433.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODI009BB12QP1&cluster=1433.r>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODI009BB12QP1. Location/Qualifiers

```
1. .1201
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CSODI009fC24"
    /tissue_type="PLACENTA COT 25-NORMALIZED"
    /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
    /notes="left strand cDNA was primed with a NotI-oligo (dr) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

FEATURES

[illegible]

RESULT 3

AL554700/c	AL554700	1201 bp	linear	EST 31-MAY-2003
LOCUS	AL554700 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI085YH07 5-PRIME, mRNA sequence.			
DEFINITION	AL554700			
ACCESSION	AL554700			
VERSION	AL554700.2	GI:31276510		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1201)			
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished (2001)			
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12895734.			

FEATURES

```

1.1201
source
accession_number
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"

```

[illegible]

```

FEATURES
source
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA003ZG04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

[illegible]

DB 904 CTGGGGYCTGTYTYYTTTGARAAAAAATAAAAAAATAAATCCCAAA 852

RESULT 6
CC317939/c
LOCUS CG317939 1304 bp DNA linear GSS 14-MAY-2003
DEFINITION TAM32-25P7 EC1.1 TAM32 Gallus gallus genomic clone TAM32-25P7,
TAM32-25P7

genomic survey sequence.

ACCESSION
CC317939
VERSION
CC317939.1 GI:30711997
KEYWORDS
GSS.
SOURCE
Gallus gallus (chicken)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE

AUTHORS
1 (bases 1 to 1304)
Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE
JOURNAL
COMMENT

Gallus gallus BAC End Reads
Unpublished (2003)
Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.edu

Insert Length: 152000 Std Error: 0.00

Seq Primer: EC1 TACGACTCACTATAGGCGG

Clas: BAC ends

High quality sequence start: 51

High quality sequence stop: 755.

FEATURES

source

1..1304
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strains="Red Jungle Fowl"
/db_xref="taxon:9031"
/clones="TAM32-25P7"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="TAM32"
/note="Vector: pCBAC1; Site 1: EcoRI; Site 2: EcoRI;
TAM32 Female Chicken library - for library and clone
ordering information: http://www.hbz.tamu.edu"

ORIGIN

Query Match 20.9%; Score 38.6; DB 28; Length 1304;
Best Local Similarity 68.8%; Pred. No. 13;
Matches 53; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 107 TGTAAAAATAGGAATAAGGGGGTTCACATTTTTCATGATATGATATATATTTT 166
DB 1101 TGTGAGAAATAGGATATAGGGGTAGGGGGAGTATTTCGATAGGTATATAATGATTT 1042
QY 167 GTATAAGAAATGGAGC 183
DB 1041 GTGAAAAGAGTGTGC 1025

RESULT 7

AW292419

LOCUS

DEFINITION
UI-H-B12-agw-h-08-0-UI-s1 NCI CGAP_Sub4 Homo sapiens cDNA clone
IMAGE:2725886 3', mRNA sequence.

ACCESSION

AW292419

VERSION

AW292419.1 GI:6699055

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A

tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html The following repetitive

elements were found in this cDNA sequence: 1-30,

>AT-rich#Low_complexity

Seq Primer: M13 Forward

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..392

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2725886"

/lab_host="DH10S (Life Technologies)"

/clone_lib="NCI CGAP Sub4"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The

NCI CGAP Sub4 library is a subtracted library derived from

the NCI CGAP Sub2 library which is a subtracted library

derived from the NCI CGAP Sub1 library, which is a

subtracted library derived from B1. B1 constitutes a

mixture of 21 normalized or subtracted NCI CGAP

libraries: NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Pr28,

NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kid5,

NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid11,

NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP CG8, NCI CGAP CLL1,

NCI CGAP Lei2, NCI CGAP Brn23, NCI CGAP Lu5,

NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP GC6,

NCI CGAP Brn25. These 21 libraries were pooled and a

single-stranded DNA preparation of the resulting mixture

was used as a tracer in a subtractive hybridization with

a driver whose composition is detailed below:

NCI CGAP Kid3 pool 1 : LAM 3334-3337, 3682-3683,

3798-3803 (IMAGE Cloneds 1323376-1323911,

1456008-1456775, 1500552-1502855) NCI CGAP Kid5 pool 1 :

LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Cloneds

1323912-1325831, 1471368-1472903, 1492104-1493255)

NCI CGAP Lu5 pool 1 : LAM 3575-3582, 3851-3854 (IMAGE

Cloneds 1414920-1417991, 1520904-1522439) NCI CGAP GC4

pool 1 : LAM 3164-3167, 3716-3720, 3733-3735 (IMAGE

Cloneds 1257096-1258631, 1469064-1470983,

1475592-1476743) NCI CGAP Pr22 pool 1 : LAM 2457-2459,

2758-2759, 3062-3068 (IMAGE Cloneds 985608-986759,

1101192-1101959, 1217928-1220615) NCI CGAP Co10 pool 1 :

LLAM 2644-2653, 2871-2872 (IMAGE Cloneds 1057416-1061255,

1144584-1145351) Subtraction was performed as previously

described (Bonaldo, Lennon & Soares (1996): Normalization

and Subtraction: Two Approaches To Facilitate Gene

Discovery. Genome Research 6, 791-806.)

TAG_TISSUE=lung

TAG_LIB=NCI CGAP_Lu5

TAG_SEQ=CAAC"

ORIGIN

Query Match 20.3%; Score 37.6; DB 10; Length 392;
Best Local Similarity 72.1%; Pred. No. 33;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 111 AAAAAATAGGAATAAGGGGGTTCACATTTTTCATGATATGATATATATTTTAT 170
DB 193 AAAGAGAGGAGGAACTAAGGTGTGCAATAATGACAGATAGTAAAAAATAATTCGTTT 252
QY 171 AAGAAAAAT 178
DB 253 AAGAAAAAT 260

RESULT 8

AW001963

LOCUS

DEFINITION

AW001963

461 bp mRNA linear EST 09-MAR-2000

Soares_thymus_NHRFth Homo sapiens cDNA clone

IMAGE:2513953 3', mRNA sequence.

```

ACCESSION      AW001963
VERSION        AW001963.1  GI:5848879
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 461)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL        Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgaps@mail.nih.gov
               This clone is available royalty-free through LLNL; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               Insert Length: 696 Std Error: 0.00
               Seq primer: 40UP from Gibco.
               High quality sequence stop: 448.

FEATURES             Location/Qualifiers
     source           1..461
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2513953"
                     /dev_stage="fetal"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="DH10B (phage-resistant)"
                     /note="Organ: thymus, pooled; Vector: pVT73D-Pac
                     (Pharmacia) with a modified polylinker; Site 1: Not I;
                     Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5',
                     TGTTCACCAATGAGTGGAGCGGCCGCAACGTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pVT73 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M. Fatima Bonaldo. "
```

Query Match 20.3%; Score 37.6; DB 9; Length 461;
Best Local Similarity 72.1%; Pred. No. 31;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Qy 111 AAAAATGAGGAATAAGGGGGTTGACATTATTTTACTGATATGATAATAATAATTGTAT 170
    |||||
Db 176 AAGAGAGGAGGAGGAACCTAAGGTGTCATTAAATGACAGATAGTAATAAATAATTGCTTT 235
    |||||
Qy 171 AAGAAAAAT 178
    |||||
Db 236 AAGAAAAAT 243
    |||||

RESULT 9
BQ000230/c
LOCUS         BQ000230.1 462 bp mRNA linear EST 17-JUN-2002
DEFINITION   UI-H-DIO-avp-j-04-0-UI.e1 NCI-CGAP_DIO Homo sapiens cDNA clone
               IMAGE:5882595 3', mRNA sequence.
ACCESSION    BQ000230
VERSION      BQ000230.1  GI:19725130
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 462)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
               Email: cgaps@mail.nih.gov
               Tissue Procurement: Dr. Jose Mercuende
```

cdNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA
sequence: 433-454, SAT rich#Low_complexity (matched complement)
Seq primer: ML3 FORWARD
POLYA=Yes.

Location/Qualifiers
1..462
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5882595"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DIO"
/note="Organ: Lung; Vector: pVT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI-CGAP DIO is a cDNA library containing the following
tissue(s): A pool of Lung Focal Fibrosis. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor digested with Not I, and cloned directionally
into pVT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ATACGGCGTC.
TAG TISSUE=Lung with fibrosis
TAG LIB=UI-H-DIO
TAG_SEQ=ATACGGCGTC"

Query Match 20.3%; Score 37.6; DB 12; Length 462;
Best Local Similarity 72.1%; Pred. No. 31;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Qy 111 AAAAATGAGGAATAAGGGGGTTGACATTATTTTACTGATATGATAATAATAATTGTAT 170
    |||||
Db 270 AAGAGAGGAGGAGGAACCTAAGGTGTCATTAAATGACAGATAGTAATAAATAATTGCTTT 211
    |||||
Qy 171 AAGAAAAAT 178
    |||||
Db 210 AAGAAAAAT 203
    |||||

RESULT 10
AI989386
LOCUS         AI989386
DEFINITION   wt80e12.x1 Soares thymus NHFT Homo sapiens cDNA clone
               IMAGE:2513806 3', mRNA sequence.
ACCESSION    AI989386
VERSION      AI989386.1  GI:5836309
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 467)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
               Email: cgaps@mail.nih.gov
               This clone is available royalty-free through LLNL; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               Seq primer: -40UP from Gibco
```

```

High quality sequence stop: 442.
Location/Qualifiers
1. 467
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2513806"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares_Chymus_MHFrh"
/notes="Organ: thymus, pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI-CGAP Lu24 was prepared, and ss circles were made in vitro. Following EAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES
source
1. 475
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2341202"
/tissue_type="carcinoid"
/lab_host="DH10B"

Query Match 20.3%; Score 37.6; DB 9; Length 467;
Best Local Similarity 72.1%; Pred. No. 31; Mismatches 0; Gaps 0;
Matches 49; Conservative 0; Indels 19;

QY 111 AAAAATAGGAATAAGGGGGTTCACATTTTCTGATATGATATATATATATATATATATAT 170
Db 176 AAAGAGAGGAGGAACCTAAGTTCATTAATGACAGATAAGTAAAAAATAATGCTTT 235

QY 171 AAGAAAT 178
Db 236 AAGAAAT 243

RESULT 11
AI697642 475 bp mRNA linear EST 18-DEC-1999
LOCUS we15902.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2341202 3',
DEFINITION mRNA sequence.
ACCESSION AI697642
VERSION AI697642
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 475)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 701 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 465.
Location/Qualifiers
1. 475
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2341202"
/tissue_type="carcinoid"
/lab_host="DH10B"

Query Match 20.3%; Score 37.6; DB 9; Length 467;
Best Local Similarity 72.1%; Pred. No. 31; Mismatches 0; Gaps 0;
Matches 49; Conservative 0; Indels 19;

QY 111 AAAAATAGGAATAAGGGGGTTCACATTTTCTGATATGATATATATATATATATATAT 170
Db 176 AAAGAGAGGAGGAACCTAAGTTCATTAATGACAGATAAGTAAAAAATAATGCTTT 235

QY 171 AAGAAAT 178
Db 236 AAGAAAT 243

RESULT 12
AI479617 564 bp mRNA linear EST 14-APR-1999
LOCUS tm32b12.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2158271 3',
DEFINITION mRNA sequence.
ACCESSION AI479617
VERSION AI479617
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 564)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 685 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 446.
Location/Qualifiers
1. 564
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2158271"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone_lib="NCI_CGAP CLL1"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATGCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

```

ORIGIN

Query Match 20.3%; Score 37.6; DB 9; Length 564;
 Best Local Similarity 72.1%; Pred. No. 29;
 Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAATAGGATRAAGGGGGTTCACATTTTACTGATGATGATATATATATTTGTAT 170
 |||||
 Db 173 AAAGAGAGGAGGAACAACTAAGGTTCATTAATGACAGATAAGTAAATAATATTCGTTT 232
 |||||

QY 171 AAGAAAAAT 178
 |||||
 Db 233 AAGAAAAAT 240
 |||||

RESULT 13
 BZ072552 699 bp DNA linear GSS 10-OCT-2002
 LOCUS lkf46509.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
 DEFINITION sequence.
 ACCESSION BZ072552
 VERSION BZ072552
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 699)
 AUTHORS Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
 Nash,W., Rabinowicz,P.D. and Wilson,R.K.
 JOURNAL Whole genome shotgun reads from Brassica oleracea
 COMMENT Unpublished (2002)
 CONTACT: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Plate: lkf46 row: b column: 09
 Seq primer: -21UPPOT forward
 Class: shotgun
 High quality sequence start: 17
 High quality sequence stop: 551.

FEATURES

source

1..699
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea002"
 /note="Vector: pOTw13; Whole genome shotgun library from
 flowering buds. DNA was purified from a crude nuclear
 prep using Brassica oleracea Tol000DH3 buds provided by
 Thomas Osborn at the University of Wisconsin. Genomic
 DNA was provided by Pablo Rabinowicz (CSHL) and the
 shotgun library prepared at Washington University Genome
 Sequencing Center."

ORIGIN

Query Match 20.3%; Score 37.6; DB 28; Length 699;
 Best Local Similarity 52.6%; Pred. No. 28;
 Matches 82; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 22 ATTGTTTGAGAAAGAACCAACCAATACCTTCTCTCTCATCAGACAGGGTATTT 81
 |||||
 Db 328 ATTGTTAAATATATAATAAATAAATAAATAATTTAATTGTAATATTATTGTTATTT 387
 |||||

QY 82 TTTATGCTCCAGACTGCCCTGTGTAAATAAATAAGAAATAAGGGGGTGTGATAT 141
 |||||
 Db 388 TTAATATTGTGAAGATCTAGATTTTTTTAGAAAACTGAGGAATACCATGTTATGTCAT 447
 |||||

QY 142 TTTACTGATGATGATATAATTTGTATAAGAAAA 177
 |||||
 Db 448 TTTATAATGAATTTAAGATTATAGATGTTATACAA 483
 |||||

RESULT 14

AU253100
 LOCUS Nb1a03504 3', mRNA sequence.
 DEFINITION

AU253100 700 bp mRNA linear EST 27-AUG-2003
 Nb1a03504 3', mRNA sequence.

ACCESSION AU253100
 VERSION AU253100.1 GI:34321813
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 700)
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Chiba, M., Morchashi, A., Inuzuka, H., Shishikura, T., Kawamoto, T.,
 Kageyama, H., Nakamura, Y., Isogai, E., Takayasu, H., Sakiyama, S.,
 Suzuki, Y., Sugano, S., Goto, T., Sato, S. and Nakagawara, A.

Expression profiling and characterization of 4200 genes cloned from
 primary neuroblastomas: identification of 305 genes differentially
 expressed between favorable and unfavorable subsets
 Oncogene 22 (35), 5525-5536 (2003)

MEDLINE 22815148
 PUBMED 12934113

COMMENT Contact: Akira Nakagawara

Division of Biochemistry
 Chiba Cancer Center Research Institute
 666-2 Nitona, Chuo-ku, Chiba 260-8717, Japan
 Tel: 81-43-264-5431(ex.5201)
 Fax: 81-43-265-4459

Email: akirana@chiba-ceri.chuo.chiba.jp.

FEATURES

source

1..700
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="Nb1a03504"
 /tissue_type="neuroblastoma"
 /clone_lib="human favorable neuroblastoma cDNA"

ORIGIN

Query Match 20.3%; Score 37.6; DB 9; Length 700;
 Best Local Similarity 72.1%; Pred. No. 28;
 Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAATAGGAATAAAGGGGGTTCACATTTTACTGATGATGATATATATATTTGTAT 170
 |||||
 Db 163 AAAGAGAGGAGGAACAACTAAGGTTCATTAATGACAGATAAGTAAATAATATTCGTTT 222
 |||||

QY 171 AAGAAAAAT 178
 |||||
 Db 223 AAGAAAAAT 230
 |||||

RESULT 15

AL543011
 LOCUS

AL543011 894 bp mRNA linear EST 12-MAY-2003
 5-PRIME, mRNA sequence.

ACCESSION AL543011
 VERSION AL543011.2 GI:30548727
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization
 Unpublished (2001)

On Feb 15, 2001 this sequence version replaced gi:12875489.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5876.f. For
more information about this cluster, see
<http://www.genoscope.cns.fr/>
<http://www.genoscope.cns.fr/cgi/seq/CS0DE013DF110P1&cluster=9876.f>. Contact :
Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE013DF11QP1.
Location/Qualifiers
1. 894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013Y122"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match	20.2%	Score	37.4;	DB	9;	Length	894;
Best Local Similarity	42.3%;	Pred.	No. 29;				
Matches	63;	Conservative	22;	Mismatches	64;	Indels	0; Gaps
QY	30	AGAAAGAAGAGACCATAAAAATACCTGCTGTGCATCAGACAGCGTATTTTTTATGCT	89				
Dd	648	AAAAAAAAAAAAAAAAAAMWTTTTTTTHMGAAMWWTTTTTTTTTTTTTTTTTTTTTKK	707				
QY	90	GTCACAGCTGCCGTGTGTAAAAATAAGGAATAAGGGGGGTGACATTATTTACTGA	149				
Dd	708	KXYDCAITTTWWWTTWWTTWWTTWWTTWAAABAAAGAGGITTTTATTTTTTTTTT	767				
QY	150	TATGTATATATTAATTTGCTATAAGAAAAT	178				
Dd	768	TTTATATTWTAAKAAAAAAAAATAAAAW	796				

Search completed: March 19, 2004, 03:16:13
Job time : 1838 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 02:14:41 ; Search time 745.5 Seconds
(without alignments)
917.966 Million cell updates/sec

Title: US-09-834-271A-27
Perfect score: 185
Sequence: 1 ggccttaaggccgcgaatc.....tgtatagaagaaatggagctc 185

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications NA:
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
 - 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
 - 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
 - 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	10	US-09-834-271A-27
2	185	100.0	185	14	US-10-406-025-6
3	183.4	99.1	185	10	US-09-834-271A-26
4	183.4	99.1	185	14	US-10-406-025-5
5	180.2	97.4	185	10	US-09-834-271A-25
6	180.2	97.4	185	14	US-10-406-025-7
7	158	85.4	2166	12	US-10-344-231-17
8	158	85.4	2166	12	US-10-363-332A-17
9	158	85.4	2267	12	US-10-344-231-18
10	158	85.4	2267	12	US-10-363-332A-18
11	158	85.4	2588	12	US-10-344-231-20
12	158	85.4	2588	12	US-10-363-332A-20
13	157.2	85.0	2604	14	US-10-146-327-5
14	154	83.2	2084	9	US-09-769-864-58
15	154	83.2	2084	9	US-09-854-346-9

16	154	83.2	2084	9	US-09-918-543-9	Sequence 9, Appli
17	154	83.2	2084	10	US-09-925-576C-9	Sequence 9, Appli
18	154	83.2	2084	12	US-10-665-667-58	Sequence 58, Appl
19	154	83.2	2084	12	US-10-644-187-3	Sequence 3, Appli
20	154	83.2	2084	14	US-10-184-771-3	Sequence 3, Appli
21	154	83.2	2084	14	US-10-186-042-3	Sequence 3, Appli
22	93.6	50.6	13222	10	US-09-736-116-60	Sequence 60, Appl
23	90.2	48.8	6837	10	US-09-928-847B-49	Sequence 49, Appl
24	44	23.8	44	10	US-09-834-271A-19	Sequence 19, Appl
25	39	21.1	42	10	US-09-834-271A-9	Sequence 9, Appli
26	38.4	20.8	43	10	US-09-834-271A-10	Sequence 10, Appl
27	37.6	20.3	2141	15	US-10-104-047-1737	Sequence 1737, Ap
28	36.4	19.7	3673778	14	US-10-312-841-2	Sequence 2, Appli
29	35.8	19.4	9121	9	US-09-070-927A-221	Sequence 221, App
30	35.6	19.2	179	9	US-09-783-590-7881	Sequence 7881, Ap
31	35.2	19.0	512	15	US-10-027-632-15736	Sequence 15736, A
32	35.2	19.0	827	15	US-10-027-632-145937	Sequence 145937,
33	35.2	19.0	19380	12	US-10-221-613-389	Sequence 389, App
34	35	18.9	11155	14	US-10-311-455-577	Sequence 577, App
35	35	18.9	14924	14	US-10-311-455-198	Sequence 198, App
36	35	18.9	14924	14	US-10-240-452-22	Sequence 22, Appli
37	34.8	18.8	3673778	14	US-10-312-841-1	Sequence 1, Appli
38	34.6	18.7	2286	12	US-10-282-122A-15549	Sequence 15549, A
39	34.6	18.7	14147	12	US-10-221-714A-470	Sequence 470, App
40	34.6	18.7	14147	14	US-10-172-086-52	Sequence 52, Appl
41	34.4	18.6	28588	9	US-09-764-887-399	Sequence 399, App
42	34.4	18.6	28588	14	US-10-073-961-399	Sequence 399, App
43	34.4	18.6	182328	12	US-10-235-192A-47	Sequence 47, Appl
44	34	18.4	510	12	US-10-085-783A-18616	Sequence 18616, A
45	34	18.4	510	15	US-10-242-535A-18616	Sequence 18616, A

ALIGNMENTS

RESULT 1

US-09-834-271A-27
; Sequence 27, Application US/09834271A
; Publication NO. US20030170876A1
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/834,271A
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US/09/258,377
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus

US-09-834-271A-27

Query Match	100.0%;	Score	185;	DB	10;	Length	185;
Best Local Similarity	100.0%;	Pred. No.	1.1e-41;				
Matches	185;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						

Qy	1	GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAAAGAACCAACCAATAAAATACCTTGT	60
Db	1	GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAAAGAACCAACCAATAAAATACCTTGT	60

Qy	61	CTGTCAATCAGACAGGGTATTTTATGCTGCTCCAGACTGCCCTGTGTAATAAATAGGA	120
Db	61	CTGTCAATCAGACAGGGTATTTTATGCTGCTCCAGACTGCCCTGTGTAATAAATAGGA	120

Qy	121	ATAAAGGGGGGTTTGACATTATTTTACTGATATGATATAATAATTTTGTATGAAGAAATGG	180
----	-----	---	-----

GENERAL INFORMATION:
APPLICANT: Widner, William
APPLICANT: Sloma, Alan
APPLICANT: Thomas, Michael D.
TITLE OF INVENTION: Methods For Producing A polypeptide In a
TITLE OF INVENTION: Bacillus Cell
FILE REFERENCE: 5455-200-US
CURRENT APPLICATION NUMBER: US/09/834,271A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: US/09/258,377
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 185
TYPE: DNA
ORGANISM: Bacillus
US-09-834-271A-25

Query Match 97.4%; Score 180.2; DB 10; Length 185;
Best Local Similarity 98.4%; Pred. No. 2.3e-40;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGCCGTGTGTAATAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGCCGTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTGGACATTTTACTGATATGTAATAATAATTTGTAAGAAATGG 180
DB 121 ATAAAGGGGGTGGTATTTTACTGATATGTAATAATAATTTGTAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 6
US-10-406-025-7
Sequence 7, Application US/10406025
Publication No. US20030186380A1
GENERAL INFORMATION:
APPLICANT: No. US20030186380A1; ozymes Biotech, Inc.
APPLICANT: Thomas, Michael D.
APPLICANT: Sloma, Alan
TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparaginase
TITLE OF INVENTION: activity
FILE REFERENCE: 10289-200-US
CURRENT APPLICATION NUMBER: US/10/406,025
CURRENT FILING DATE: 2003-04-01
PRIOR APPLICATION NUMBER: US 60/369,192
PRIOR FILING DATE: 2002-04-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.2
SEQ ID NO 7
LENGTH: 185
TYPE: DNA
ORGANISM: Bacillus
US-10-406-025-7

Query Match 97.4%; Score 180.2; DB 14; Length 185;
Best Local Similarity 98.4%; Pred. No. 2.3e-40;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAAATAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAAATAATAGGA 120

QY 121 ATAAAGGGGGTGGACATTTTACTGATATGTAATAATAATTTGTAAGAAATGG 180
DB 121 ATAAAGGGGGTGGTATTTTACTGATATGTAATAATAATTTGTAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 7
US-10-344-231-17
Sequence 17, Application US/10344231
Publication No. US20040038845A1
GENERAL INFORMATION:
APPLICANT: Pedersen, Poul
APPLICANT: Roggen, Erwin Ludo
TITLE OF INVENTION: Improved production of proteases with inhibitors
FILE REFERENCE: 10064-204-US
CURRENT APPLICATION NUMBER: US/10/344,231
CURRENT FILING DATE: 2003-02-10
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.2
SEQ ID NO 17
LENGTH: 2166
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-344-231-17

Query Match 85.4%; Score 158; DB 12; Length 2166;
Best Local Similarity 97.0%; Pred. No. 8.4e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTGTGTATCAGACA 73
DB 172 TGCATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTGTGTATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGTGTGTAAATAATAGGAATAGGGGGGT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGTGTGTAAATAATAGGAATAGGGGGGT 291

QY 134 GACATTTTACTGATATGTAATAATAATTTGTAAGAAATG 179
DB 292 GTTATTATTACTGATATGTAATAATAATTTGTAAGAAATG 337

RESULT 8
US-10-363-332A-17
Sequence 17, Application US/10363332A
Publication No. US20040038375A1
GENERAL INFORMATION:
APPLICANT: Pedersen, Poul
APPLICANT: No. US20040038375A1; regaard-Madsen, Mads
TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
FILE REFERENCE: 10080-204-US
CURRENT APPLICATION NUMBER: US/10/363,332A
CURRENT FILING DATE: 2003-05-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.2
SEQ ID NO 17
LENGTH: 2166
TYPE: DNA
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Construct A
US-10-363-332A-17

Query Match 85.4%; Score 158; DB 12; Length 2166;

Best Local Similarity 97.0%; Score 158; DB 12; Length 2267;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 291

QY 134 GACATTATTTTACTGATGTATATAATAATTTGTATAAGAAAATG 179
DB 292 GTTATTTTACTGATGTATAATAATTTGTATAAGAAAATG 337

RESULT 9
US-10-344-231-18
; Sequence 18, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-344-231-18

Query Match 85.4%; Score 158; DB 12; Length 2267;
Best Local Similarity 97.0%; Pred. No. 8.5e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 291

QY 134 GACATTATTTTACTGATGTATATAATAATTTGTATAAGAAAATG 179
DB 292 GTTATTTTACTGATGTATAATAATTTGTATAAGAAAATG 337

RESULT 10
US-10-363-332A-18
; Sequence 18, Application US/10363332A
; Publication No. US20040038375A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: No. US20040038375A1regaard-Madsen, Mads
; TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
; FILE REFERENCE: 10080.204-US
; CURRENT APPLICATION NUMBER: US/10/363,332A
; CURRENT FILING DATE: 2003-05-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Construct B

US-10-363-332A-18

Query Match 85.4%; Score 158; DB 12; Length 2267;
Best Local Similarity 97.0%; Pred. No. 8.5e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 291

QY 134 GACATTATTTTACTGATGTATATAATAATTTGTATAAGAAAATG 179
DB 292 GTTATTTTACTGATGTATAATAATTTGTATAAGAAAATG 337

RESULT 11
US-10-344-231-20
; Sequence 20, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 2588
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2070)..(2070)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2077)..(2077)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2084)..(2084)
; OTHER INFORMATION: n is a, c, g, or t
US-10-344-231-20

Query Match 85.4%; Score 158; DB 12; Length 2588;
Best Local Similarity 97.0%; Pred. No. 9e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 291

QY 134 GACATTATTTTACTGATGTATATAATAATTTGTATAAGAAAATG 179
DB 292 GTTATTTTACTGATGTATAATAATTTGTATAAGAAAATG 337

RESULT 12
US-10-363-332A-20
; Sequence 20, Application US/10363332A
; Publication No. US20040038375A1

GENERAL INFORMATION:
 APPLICANT: Pedersen, Poul
 APPLICANT: No. US20040039375Alregaard-Madsen, Mads
 TITLE OF INVENTION: Method for Screening Highly Active Proteases and Inhibitors
 FILE REFERENCE: 10080.204-US
 CURRENT APPLICATION NUMBER: US/10/363.332A
 CURRENT FILING DATE: 2003-05-27
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 20
 LENGTH: 2588
 TYPE: DNA
 ORGANISM: Artificial
 FEATURE:
 OTHER INFORMATION: Construct D
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (2070)..(2070)
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (2077)..(2077)
 OTHER INFORMATION: n is a, c, g, or t
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (2084)..(2084)
 OTHER INFORMATION: n is a, c, g, or t
 US-10-363-332A-20

Query Match 85.4%; Score 158; DB 12; Length 2588;
 Best Local Similarity 97.0%; Pred. No. 9e-34; Mismatches 5; Indels 0; Gaps 0;
 Matches 161; Conservative 0;

Qy 14 TCCAATCGATTGTTTGAGAAAAGAGAGACCATAAAAAATACCTTGTCTGTCATCAGACA 73
 Db 172 TCCAATCGATTGTTTGAGAAAAGAGAGACCATAAAAAATACCTTGTCTGTCATCAGACA 231

Qy 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAATAGGAATAAGGGGGTT 133
 Db 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAATAGGAATAAGGGGGTT 291

Qy 134 GACATATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 179
 Db 292 GTATATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 337

RESULT 13
 US-10-146-327-5
 Sequence 5, Application US/10146327
 Publication No. US2003004954A1
 GENERAL INFORMATION:
 APPLICANT: Andersen, Carsten
 APPLICANT: Jorgensen, Christel T.
 APPLICANT: Bisgard-Frantzen, Henrik
 APPLICANT: Svendsen, Allan
 APPLICANT: Kjaerulff, Soren
 TITLE OF INVENTION: Alpha-Amylase Variants
 FILE REFERENCE: 5886.200-US
 CURRENT APPLICATION NUMBER: US/10/146,327
 CURRENT FILING DATE: 2002-05-15
 PRIOR APPLICATION NUMBER: US/09/537,168
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: PA 1999 00437
 PRIOR FILING DATE: 1999-03-30
 PRIOR APPLICATION NUMBER: 60/127,427
 PRIOR FILING DATE: 1999-04-01
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 5
 LENGTH: 2604
 TYPE: DNA
 ORGANISM: Bacillus amyloliquefaciens
 US-10-146-327-5

Query Match 85.0%; Score 157.2; DB 14; Length 2604;
 Best Local Similarity 98.1%; Pred. No. 1.5e-33;
 Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAAGAGAGACCATAAAAAATACCTTGTCTGTCATCAGACAGGGT 77
 Db 73 ATCGATTGTTTGAGAAAAGAGAGACCATAAAAAATACCTTGTCTGTCATCAGACAGGGT 132

Qy 78 ATTTTATGCTGCCAGACTGCCGCTGTGTAATAATAGGAATAAGGGGGTTGACA 137
 Db 133 ATTTTATGCTGCCAGACTGCCGCTGTGTAATAATAGGAATAAGGGGGTTGTTA 192

Qy 138 TTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 179
 Db 193 TTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 234

RESULT 15
 US-09-854-346-9
 Sequence 9, Application US/09854346
 Patent No. US20020068352A1
 GENERAL INFORMATION:
 APPLICANT: No. US20020068352A1ozymes A/S
 APPLICANT: Svendsen, Allan
 APPLICANT: Jorgensen, Christel Thea
 APPLICANT: Nielsen, Bjarne Ronfeldt
 TITLE OF INVENTION: Alpha-amylase variants with altered 1.6 activity

```

; FILE REFERENCE: 6140.200-US
; CURRENT APPLICATION NUMBER: US/09/854,346
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)..(1794)
; OTHER INFORMATION: BAN
US-09-834-346-9

Query Match      83.2%; Score 154; DB 9; Length 2084;
Best Local Similarity 96.9%; Pred.No.1.1e-32;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCATCAGACAGGGT 77
Db 73 ATCGATTGTTTGAGAAAGAGAGAGACCATATAAAATACCTTGTCTGTCATCAGACAGGGT 132
Qy 78 ATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAATAGGAATAAAGGGGGTTGACA 137
Db 133 ATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAATAGGAATAAAGGGGGTTGTTA 192
Qy 138 TTATTTACTGATATGATATATATAATTTGTATAGAAATG 179
Db 193 TTATTTACTGATATGATATATAATTTGTATAGAAATG 234

```

Search completed: March 19, 2004, 05:16:24
Job time : 750.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 00:21:10 ; Search time 215 Seconds
(without alignments)
3655.426 Million cell updates/sec

Title: US-09-834-271A-27
Perfect score: 185
Sequence: 1 ggccttagggcgtgcaatc.....tgatagaagaatggagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	2	Aaz23324 B. amylol
2	185	100.0	185	4	Aaf62647 Consensus
3	185	100.0	185	4	Aad09924 Bacillus
4	183.4	99.1	185	2	Aaz23323 B. amylol
5	183.4	99.1	185	4	Aaf62646 Consensus
6	183.4	99.1	185	4	Aad09923 Bacillus
7	180.2	97.4	185	2	Aaz23322 B. thurin
8	180.2	97.4	185	4	Aaf62645 Wild type
9	180.2	97.4	185	4	Aad09922 Bacillus
10	178.6	96.5	185	2	Aaz23325 Bacillus
11	178.6	96.5	185	4	Aad09931 Bacillus
12	158	85.4	2166	6	Aad34433 C-termina
13	158	85.4	2166	6	ABL40498 DNA const
14	158	85.4	2267	6	Aad34434 Subtilisi
15	158	85.4	2267	6	ABL40499 DNA const
16	158	85.4	2588	6	Aad34436 PamyL ATG
17	158	85.4	2588	6	ABL40501 DNA const
18	157.2	85.0	249	3	Aaa93677 Bacillus
19	157.2	85.0	249	7	ACA61514 Modified
20	157.2	85.0	270	3	AAA93678 Modified
21	157.2	85.0	2604	4	Aaz21079 Bacillus
22	157.2	85.0	2604	4	AAA37850 B. amylol
23	154.8	83.7	10216	2	Aat39279 Transposo

ALIGNMENTS

RESULT 1
Aaz23324
ID Aaz23324 standard; DNA; 185 BP.

XX AC Aaz23324;
XX DT 06-DEC-1999 (first entry)
XX DE B. amyloliquefaciens amyQ promoter DNA #2.

XX DE Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
XX KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
XX OS Bacillus amyloliquefaciens.

XX PN WO9943835-A2.
XX PD 02-SEP-1999.

XX PF 26-FEB-1999; 99WO-US004360.

XX PR 26-FEB-1998; 98US-00031442.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PI Widner W, Sioma A, Thomas MD;

XX DR WPI; 1999-561370/47.

XX PT Production of polypeptide in Bacillus using specific promoters,
XX PT particularly for producing enzymes.

XX PS Claim 44; Page 90; 90pp; English.

XX CC This invention describes a novel method for the production of a
XX CC polypeptide in Bacillus using specific tandem or consensus promoters. The
XX CC method is used to produce homologous or particularly heterologous
XX CC proteins, particularly enzymes (specifically serine protease, maltogenic
XX CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
XX CC etc. The specified promoters provide increased expression of the sequence
XX CC which encodes the polypeptide of the invention. After incorporation of
XX CC the nucleic acid construct of the invention, any marker gene may be
XX CC deleted, resulting in a cell that is preferred for environmental and
XX CC regulatory regions. This sequence represents a Bacillus amyloliquefaciens
XX CC alpha-amylase amyQ promoter which is used to describe the method of the
XX CC invention

24	154	83.2	162	2	AAT85631
25	154	83.2	537	2	AAQ22579
26	154	83.2	576	1	AAZ20046
27	154	83.2	2084	2	AAQ8067
28	154	83.2	2084	2	AAQ95032
29	154	83.2	2084	2	AAQ95032
30	154	83.2	2084	2	AAQ95032
31	154	83.2	2084	3	AAQ95032
32	154	83.2	2084	6	AAQ95032
33	154	83.2	2084	6	AAQ95032
34	154	83.2	2084	6	AAQ95032
35	154	83.2	2084	6	AAQ95032
36	154	83.2	2084	6	AAQ95032
37	154	83.2	2084	6	AAQ95032
38	154	83.2	2084	6	AAQ95032
39	154	83.2	2084	6	AAQ95032
40	154	83.2	2084	6	AAQ95032
41	154	83.2	2084	6	AAQ95032
42	154	83.2	2084	6	AAQ95032
43	154	83.2	2084	6	AAQ95032
44	154	83.2	2084	6	AAQ95032
45	154	83.2	2084	6	AAQ95032

CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
 CC subillisin Carlsberg gene promoter have the sequence TTGACA for the -35
 CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a Bacillus strain. The present
 CC sequence is Bacillus amyloliquefaciens alpha-amylase (amyQ) consensus
 CC mutated promoter, used in the exemplification of the invention
 XX
 SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;

Query Match 100.0%; Score 185; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 3.1e-42; Indels 0; Gaps 0;
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGT 60
 DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGT 60
 QY 61 CTGTCTATCAGACAGGGTATTTTTTATGCTGCTCCAGACTGCGCTGTGTAATAATAGGA 120
 DB 61 CTGTCTATCAGACAGGGTATTTTTTATGCTGCTCCAGACTGCGCTGTGTAATAATAGGA 120
 QY 121 ATAAAGGGGGTGTGACATTATTTTACTGATATGATATATATAATTTGTATAGAAATGG 180
 DB 121 ATAAAGGGGGTGTGACATTATTTTACTGATATGATATATATAATTTGTATAGAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 4
 AAZ23323
 ID AAZ23323 standard; DNA; 185 BP.
 AC AAZ23323;
 DT 06-DEC-1999 (first entry)
 DE B. amyloliquefaciens amyQ promoter DNA #1.
 XX Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
 KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
 XX Bacillus amyloliquefaciens.
 XX WO943835-A2.
 XX 02-SEP-1999.
 XX 26-FEB-1999; 99WO-US004360.
 XX 26-FEB-1998; 98US-00031442.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX Widner W, Sloma A, Thomas MD;
 XX WPI; 1999-561370/47.
 XX Production of polypeptide in Bacillus using specific promoters,
 PT particularly for producing enzymes.
 XX Claim 44; Page 90; 90pp; English.
 XX This invention describes a novel method for the production of a
 CC polypeptide in Bacillus using specific tandem or consensus promoters. The
 CC method is used to produce homologous or particularly heterologous
 CC proteins, particularly enzymes (specifically serine protease, maltogenic
 CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
 CC etc. The specified promoters provide increased expression of the sequence
 CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be

CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents a Bacillus amyloliquefaciens
 CC alpha-amylase amyQ promoter which is used to describe the method of the
 CC invention
 XX
 SQ Sequence 185 BP; 63 A; 25 C; 42 G; 55 T; 0 U; 0 Other;

Query Match 99.1%; Score 183.4; DB 2; Length 185;
 Best Local Similarity 99.5%; Pred. No. 8.8e-42; Indels 0; Gaps 0;
 Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGT 60
 DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGT 60
 QY 61 CTGTCTATCAGACAGGGTATTTTTTATGCTGCTCCAGACTGCGCTGTGTAATAATAGGA 120
 DB 61 CTGTCTATCAGACAGGGTATTTTTTATGCTGCTCCAGACTGCGCTGTGTAATAATAGGA 120
 QY 121 ATAAAGGGGGTGTGACATTATTTTACTGATATGATATATAATTTGTATAGAAATGG 180
 DB 121 ATAAAGGGGGTGTGACATTATTTTACTGATATGATATATAATTTGTATAGAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 5
 AAF62646
 ID AAF62646 standard; DNA; 185 BP.
 AC AAF62646;
 DT 03-MAY-2001 (first entry)
 DE Consensus amyQ promoter.
 XX Pectin acetylase; degrade; plant cell wall; ss.
 XX Unidentified.
 XX US6184028-B1.
 XX 06-FEB-2001.
 XX 26-AUG-1999; 99US-00384305.
 XX 26-AUG-1999; 99US-00384305.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX Thomas MD, Brown KM;
 XX WPI; 2001-190946/19.
 XX Novel isolated polypeptide having pectin acetylase activity useful
 PT for degrading pectic substances and in degradation or modification of
 PT acetylated pectins and plant cell walls.
 XX Example 9; Col 43-44; 35pp; English.
 XX The present invention relates to Bacillus subtilis pectin acetylase
 CC protein. The invention is useful for degrading a pectic substance. It is
 CC also useful for degrading soluble and insoluble pectins with varying
 CC degrees of esterification, clarification etc. The protein may be used
 CC alone or in combination with other enzymes for the degradation or
 CC modification of acetylated pectins, degradation or modification of plant
 CC cell walls
 XX Sequence 185 BP; 63 A; 25 C; 42 G; 55 T; 0 U; 0 Other;
 SQ Query Match 99.1%; Score 183.4; DB 4; Length 185;

CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents a *Bacillus thuringiensis*
CC alpha-amylase amyQ promoter which is used to describe the method of the
CC invention
XX
SQ Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;

Query Match 97.4%; Score 180.2; DB 2; Length 185;
Best Local Similarity 98.4%; Pred. No. 6.9e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTCAATCGATTGTTTGAGAAAGAGAGACCATTAATAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTCAATCGATTGTTTGAGAAAGAGAGACCATTAATAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
DB 61 CTGTCATCAGACAGAGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
QY 121 ATAAAGGGGGGTGACATTTTACTGATATGTAATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGGTGTTTATTATTTTACTGATATGTAATAATTTGTATAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 8

AAF62645
ID AAF62645 standard; DNA; 185 BP.

XX AAF62645;

XX 03-MAY-2001 (first entry)

XX Wild type amyQ promoter.

XX Pectin acetyltransferase; degrade; plant cell wall; ss.

XX Unidentified.

XX US6184028-B1.

XX 06-FEB-2001.

XX 26-AUG-1999; 99US-00384305.

XX 26-AUG-1999; 99US-00384305.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Thomas MD, Brown KM;

XX WPI; 2001-190946/19.

XX Novel isolated polypeptide having pectin acetyltransferase activity useful
XX for degrading pectic substances and in degradation or modification of
XX acetylated pectins and plant cell walls.

XX Example 9; Fig 9; 35pp; English.

XX The present invention relates to *Bacillus subtilis* pectin acetyltransferase
XX protein. The invention is useful for degrading a pectic substance. It is
XX also useful for degrading soluble and insoluble pectins with varying
XX degrees of esterification, clarification etc. The protein may be used
XX alone or in combination with other enzymes for the degradation or
XX modification of acetylated pectins, degradation or modification of plant
XX cell walls

XX Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;

Query Match 97.4%; Score 180.2; DB 4; Length 185;
Best Local Similarity 98.4%; Pred. No. 6.9e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTCAATCGATTGTTTGAGAAAGAGAGACCATTAATAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTCAATCGATTGTTTGAGAAAGAGAGACCATTAATAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
DB 61 CTGTCATCAGACAGAGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
QY 121 ATAAAGGGGGGTGACATTTTACTGATATGTAATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGGTGTTTATTATTTTACTGATATGTAATAATTTGTATAAGAAATGG 180

QY 181 AGCTC 185

DB 181 AGCTC 185

RESULT 9

AAD09922
ID AAD09922 standard; DNA; 185 BP.

XX AAD09922;

XX 12-SEP-2001 (first entry)

XX *Bacillus amyloliquefaciens* alpha-amylase (amyQ) promoter.

XX *Bacillus* cell; tandem promoter; consensus promoter;

XX polypeptide production; amyL promoter; amyQ promoter; aprH promoter;

XX *cryIIIA* promoter; *subtilisin* Carlsberg gene promoter; alpha-amylase; ds.

XX *Bacillus amyloliquefaciens*.

XX US6255076-B1.

XX 03-JUL-2001.

XX 26-FEB-1999; 99US-00258377.

XX 26-FEB-1998; 98US-00031442.

XX (NOVO) NOVOZYMES BIOTECH INC.

XX Widner W, Sloma A, Thomas MD;

XX WPI; 2001-440518/47.

XX Producing a polypeptide in a *Bacillus* strain comprises cultivating a
XX *Bacillus* cell, which contains a nucleic acid construct comprising a
XX tandem promoter and consensus promoters.

XX Example 18; Col 43-44; 54pp; English.

XX The present invention relates to a method for producing a polypeptide,
XX comprises cultivating a *Bacillus* cell, which contains a nucleic acid
XX construct comprising a tandem promoter and consensus promoters. The
XX *Bacillus* is cultivated in a medium conducive for the production of the
XX polypeptide. Each promoter sequence of the tandem promoter is operably
XX linked to a nucleic acid sequence encoding the polypeptide. The tandem
XX promoter comprises amyL promoter, amyQ promoter, aprH promoter, *cryIIIA*
XX promoter or *subtilisin* Carlsberg gene promoter. The consensus promoters
XX of the amyL promoter, amyQ promoter, aprH promoter, *cryIIIA* promoter or
XX *subtilisin* Carlsberg gene promoter have the sequence TTGACA for the -35
XX region and TATAAT for the -10 region. The method further comprises
XX isolating the polypeptide from the cultivation medium. The method is
XX useful for producing a polypeptide in a *Bacillus* strain. The present
XX sequence is *Bacillus amyloliquefaciens* alpha-amylase (amyQ) promoter,
XX used in the exemplification of the invention. Note: This sequence is

CC stated as being the same as that shown as SEQ ID NO 25 (AAD09931) in
CC figure 21 of the specification. However the sequences differ by at
CC position 116
XX
SQ Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;
Query Match 97.4%; Score 180.2; DB 4; Length 185;
Best Local Similarity 98.4%; Pred. No. 6.9e-41; Indels 0; Gaps 0;
Matches 182; Conservative 0; Mismatches 3;
QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTAATTTTATGCTGCCAGACTGCCGTGTGTAATAAATAGGA 120
DB 61 CTGTCATCAGACAGAGGTAATTTTATGCTGCCAGACTGCCGTGTGTAATAAATAGGA 120
QY 121 ATAAAGGGGGTGTGACATATTTTACTGATATGTAATAATAATTTCTATAAGAAAATGG 180
DB 121 ATAAAGGGGGTGTGTAATTTTACTGATATGTAATAATAATTTCTATAAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185
RESULT 10
AAZ23325
ID AAZ23325 standard; DNA; 185 BP.
XX
AC AAZ23325;
XX
DT 06-DEC-1999 (first entry)
XX
DE Bacillus sp. amyQ promoter consensus DNA.
XX
XX Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
XX antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
OS Bacillus sp.
XX
PN WO9943835-A2.
XX
XX 02-SEP-1999.
PD
XX 26-FEB-1999; 99WO-US004360.
PF
XX 26-FEB-1998; 98US-00031442.
PR
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX
XX Widner W, Sloma A, Thomas MD;
PI
XX WPI; 1999-561370/47.
DR
XX Production of polypeptide in Bacillus using specific promoters,
PT particularly for producing enzymes.
XX
XX Disclosure; Fig 21; 90pp; English.
PS
XX This invention describes a novel method for the production of a
CC polypeptide in Bacillus using specific tandem or consensus promoters. The
CC method is used to produce homologous or particularly heterologous
CC proteins, particularly enzymes (specifically serine protease, maltogenic
CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
CC etc. The specified promoters provide increased expression of the sequence
CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents the consensus sequence of a
CC Bacillus sp. alpha-amylase amyQ promoter which is used to describe the
CC method of the invention

XX
SQ Sequence 185 BP; 63 A; 24 C; 42 G; 56 T; 0 U; 0 Other;
Query Match 96.5%; Score 178.6; DB 2; Length 185;
Best Local Similarity 97.8%; Pred. No. 1.9e-40;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTAATTTTATGCTGCCAGACTGCCGTGTGTAATAAATAGGA 120
DB 61 CTGTCATCAGACAGAGGTAATTTTATGCTGCCAGACTGCCGTGTGTAATAAATAGGA 120
QY 121 ATAAAGGGGGTGTGACATATTTTACTGATATGTAATAATAATTTGTAAGAAAATGG 180
DB 121 ATAAAGGGGGTGTGTAATTTTACTGATATGTAATAATAATTTGTAAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185
RESULT 11
AAD09931
ID AAD09931 standard; DNA; 185 BP.
XX
AC AAD09931;
XX
DT 12-SEP-2001 (first entry)
XX
DE Bacillus amyloliquefaciens amyQ promoter, alternative version.
XX
XX Bacillus cell; tandem promoter; consensus promoter;
KW polypeptide production; amyL promoter; amyQ promoter; apxH promoter;
KW cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.
XX
OS Bacillus amyloliquefaciens.
XX
FH Key Location/Qualifiers
FT misc_feature 116
FT /tag= a
FT /note= "Base T is found at this location in the sequence
FT shown in the sequence listing of the specification
FT (AAD09922)"
XX
XX US6255076-B1.
XX PN
XX PD 03-JUL-2001.
XX
XX 26-FEB-1999; 99US-00258377.
PF
XX 26-FEB-1998; 98US-00031442.
PR
XX (NOVO) NOVOZYMES BIOTECH INC.
XX PA
XX Widner W, Sloma A, Thomas MD;
PI
XX WPI; 2001-440518/47.
DR
XX Producing a polypeptide in a Bacillus strain comprises cultivating a
PT Bacillus cell, which contains a nucleic acid construct comprising a
PT tandem promoter and consensus promoters.
XX
XX Example 18; Fig 21; 54pp; English.
PS
XX The present invention relates to a method for producing a polypeptide,
CC comprises cultivating a Bacillus cell, which contains a nucleic acid
CC construct comprising a tandem promoter and consensus promoters. The
CC Bacillus is cultivated in a medium conducive for the production of the
CC polypeptide. Each promoter sequence of the tandem promoter is operably
CC linked to a nucleic acid sequence encoding the polypeptide. The tandem

CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
 CC promoter or subtilisin carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
 CC subtilisin carlsberg gene promoter have the sequence TTGACA for the -35
 CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a Bacillus strain. The present
 CC sequence is an alternative of Bacillus amyloliquefaciens alpha-amylase
 CC (amyQ) promoter, used in the exemplification of the invention. Note: This
 CC sequence is stated as being the same as that shown as SEQ ID NO 25
 CC (AAD09922) in column 43-44 of the specification. However the sequences
 CC differ by at position 116
 XX
 SQ Sequence 185 BP; 63 A; 24 C; 42 G; 56 T; 0 U; 0 Other;

Query Match 96.5%; Score 178.6; DB 4; Length 185;
 Best Local Similarity 97.8%; Pred. No. 1.9e-40;
 Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GGCTTTAAGGGCTGCAATCGATTGTTTCAGAAAAGAGAACACCAATAAATACCTTGT 60
 DB 1 GACCTTAAGGGCTGCAATCGATTGTTTCAGAAAAGAGAACACCAATAAATACCTTGT 60
 QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAATAAGGA 120
 DB 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAATAAGGA 120
 QY 121 ATAAAGGGGGTGCACATATTTTCTGATATGATATATATATATATATATATATATATAT 180
 DB 121 ATAAAGGGGGTGTATTTTATTTTCTGATATGTAATGTAATATATATATATATATAT 180

QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 12

AAD34433
 ID AAD34433 standard; DNA; 2166 BP.

XX AC AAD34433;

XX DT 29-AUG-2003 (revised)

XX DT 16-JUL-2002 (first entry)

XX DE C-terminal Subtilisin 309-CI-2A fusion protein encoding construct A DNA.

XX KW Protease; screening; barley; chymotrypsin inhibitor; CI-2A; laundry;
 XX KW detergent industry; dishwash; fabric softener; subtilisin 309; savinase;
 XX KW chimeric; ds.

XX OS Hordeum sp.

XX OS Bacillus lentus.

XX OS Chimeric.

XX PN W0200218588-A1.

XX XX 07-MAR-2002.

XX XX 17-JUL-2001; 2001WO-DK000503.

XX XX 29-AUG-2000; 2000DK-00001273.

XX XX 05-SEP-2000; 2000US-0230017P.

XX XX (NOVO) NOVOZYMES AS.

XX XX Pedersen PE, Norregaard-Madsen M;

XX XX WPI; 2002-329776/36.

XX XX Screening protease/protease inhibitor gene library for gene encoding the
 PT same, by constructing host cell having the gene, cultivating cell,
 PT dissociating complex formed between them, selecting protein and isolating

PT gene.

XX Example 1; Page 47-48; 54pp; English.

XX The invention relates to a method for screening a protease/protease
 CC inhibitor gene library for a gene encoding a protease of interest. The
 CC method comprises introducing a first gene of the protease gene library
 CC and a second gene encoding a protease inhibitor into a host cell;
 CC cultivating the host cell, wherein the cell expresses the first and the
 CC second genes to produce a complex of a protease and the inhibitor;
 CC dissociating the inhibitor from the complex and selecting the protease of
 CC interest and isolating the encoding gene. A complex may be used for a
 CC number of industrial applications, in particular within the detergent
 CC industry, thus, relates to a cleaning or detergent composition,
 CC preferably a laundry or dishwash composition comprising the complex,
 CC where hand or machine laundry detergent composition including a laundry
 CC additive composition suitable for pre-treatment of stained fabrics and a
 CC rinse added composition, or be formulated as a detergent composition for
 CC use in general household hard surface cleaning operations or be
 CC formulated for hand or machine dish washing operations. The present
 CC sequence is a construct encoding a fusion protein containing C-terminal
 CC region of subtilisin 309 (savinase) and barley chymotrypsin inhibitor CI-
 CC 2A protein used in the invention. (Updated on 29-AUG-2003 to standardise
 CC OS field)

SQ Sequence 2166 BP; 610 A; 442 C; 582 G; 532 T; 0 U; 0 Other;

Query Match 85.4%; Score 158; DB 6; Length 2166;

Best Local Similarity 97.0%; Pred. No. 2e-34;

Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCAATCGATTGTTTGAGAAAAGAGAACCAATAAATACCTTGTCTCTCATCAGACA 73

DB 172 TGCAATCGATTGTTTGAGAAAAGAGAACCAATAAATACCTTGTCTCTCATCAGACA 231

QY 74 GGGTATTTTTTATGCTGTCAGACTGTCGCTGTGTAAAAATAGGAATAAGGGGGTT 133

DB 232 GGGTATTTTTTATGCTGTCAGACTGTCGCTGTGTAAAAATAGGAATAAGGGGGTT 291

QY 134 GACATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 179

DB 292 GTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 337

RESULT 13

ABL40498

ID ABL40498 standard; DNA; 2166 BP.

XX AC ABL40498;

XX DT 17-JUN-2002 (first entry)

XX DE DNA construct A comprising Savinase-CI-2A fusion sequence.

XX KW Protease; allergenicity; detergent; subtilase; S11; S12; additive;
 XX KW protease inhibitor; barley; chymotrypsin; CI-2A; fusion protein;
 XX KW subtilisin; savinase; apr gene; ds.

XX OS Synthetic.

XX OS Bacillus sp.

XX OS Hordeum sp.

XX PN W0200216619-A1.

XX XX 28-FEB-2002.

XX XX 09-JUL-2001; 2001WO-DK000479.

XX XX 21-AUG-2000; 2000DK-00001233.

XX XX 24-AUG-2000; 2000US-0227501P.

XX XX (NOVO) NOVOZYMES AS.

```

PI Pedersen PE;
XX
XX WPI; 2002-242031/29.
XX
XX Producing protease-inhibitor compositions with reduced allergenicity,
XX stability and activity for use in detergent compositions.
XX
XX Example 1; Page 45-46; 52pp; English.
XX
XX The invention provides a method for producing a protease-inhibitor
XX complex with reduced allergenicity for use in detergents. The method
XX involves (a) constructing a fusion polynucleotide sequence in frame, the
XX sequence comprising a first gene encoding a protease (preferably a
XX subtilase S11 or S12) and a second gene encoding a protease inhibitor;
XX (b) introducing the sequence into a host cell; and (c) cultivating the
XX host cell (the host cell expresses the sequence and produces a non-
XX covalently linked complex of the protease and the inhibitor). The
XX protease-inhibitor complex produced by the method may be used as an
XX nucleotide sequence of a DNA construct comprising the savinase-CI-2A
XX fusion sequence
XX
XX Sequence 2166 BP; 610 A; 442 C; 582 G; 532 T; 0 U; 0 Other;
XX
XX Query Match 85.4%; Score 158; DB 6; Length 2166;
XX Best Local Similarity 97.0%; Pred. No. 2e-34;
XX Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX QY 14 TGCATCGATTGTTTGAGAAAGAGAGAGACCATAAAATACCTTGTCTCATCAGACA 73
XX |
XX Db 172 TGCATCGATTGTTTGAGAAAGAGAGAGACCATAAAATACCTTGTCTCATCAGACA 231
XX |
XX QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAATAGGAATAAGGGGGTT 133
XX |
XX Db 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAATAGGAATAAGGGGGTT 291
XX |
XX QY 134 GACATTATTTACTGATATGTAATAATAATTTGTATAAGAAATG 179
XX |
XX Db 292 GTTATTATTTACTGATATGTAATAATAATTTGTATAAGAAATG 337
XX |
XX
XX RESULT 14
XX AAD34434
XX ID AAD34434 standard; DNA; 2267 BP.
XX
XX AC AAD34434;
XX
XX DT 29-AUG-2003 (revised)
XX DT 16-JUL-2002 (first entry)
XX
XX DE Subtilisin 309-CI-2A transcriptional product encoding construct B DNA.
XX
XX KW Protease; screening; barley; chymotrypsin inhibitor; CI-2A; laundry;
XX detergent industry; dishwash; fabric softener; subtilisin 309; savinase;
XX chimeric; ds.
XX
XX OS Hordeum sp.
XX OS Bacillus lentus.
XX OS Chimeric.
XX
XX PN WO200216588-A1.
XX
XX PD 07-MAR-2002.
XX
XX PF 17-JUL-2001; 2001WO-DK000503.
XX
XX PR 29-AUG-2000; 2000DK-00001273.
XX PR 05-SEP-2000; 2000US-0230017P.
XX
XX PA (NOVO ) NOVOZYMES AS.
XX
XX PI Pedersen PE, Norregaard-Madsen M;
XX WPI; 2002-329776/36.
XX
XX
XX Screening protease/protease inhibitor gene library for gene encoding the
XX same, by constructing host cell having the gene, cultivating cell,
XX dissociating complex formed between them, selecting protein and isolating
XX gene.
XX
XX Example 1; Page 48; 54pp; English.
XX
XX The invention relates to a method for screening a protease/protease
XX inhibitor gene library for a gene encoding a protease of interest. The
XX method comprises introducing a first gene of the protease gene library
XX and a second gene encoding a protease inhibitor into a host cell;
XX cultivating the host cell, wherein the cell expresses the first and the
XX second genes to produce a complex of a protease and the inhibitor;
XX dissociating the inhibitor from the complex and selecting the protease of
XX interest and isolating the encoding gene. A complex may be used for a
XX number of industrial applications, in particular within the detergent
XX industry, thus, relates to a cleaning or detergent composition,
XX preferably a laundry or dishwash composition comprising the complex,
XX where hand or machine laundry detergent composition including a laundry
XX additive composition suitable for pre-treatment of stained fabrics and a
XX rinse added composition, or be formulated as a detergent composition for
XX use in general household hard surface cleaning operations or be
XX formulated for hand or machine dish washing operations. The product
XX sequence is a construct encoding a transcriptional fusion product
XX containing C-terminal region of subtilisin 309 (savinase) and barley
XX chymotrypsin inhibitor CI-2A protein used in the invention. (Updated on
XX 29-AUG-2003 to standardise CS field)
XX
XX SQ Sequence 2267 BP; 640 A; 462 C; 601 G; 564 T; 0 U; 0 Other;
XX
XX Query Match 85.4%; Score 158; DB 6; Length 2267;
XX Best Local Similarity 97.0%; Pred. No. 2e-34;
XX Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATAAAATACCTTGTCTCATCAGACA 73
XX |
XX Db 172 TGCATCGATTGTTTGAGAAAGAGAGACCATAAAATACCTTGTCTCATCAGACA 231
XX |
XX QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAATAGGAATAAGGGGGTT 133
XX |
XX Db 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAATAGGAATAAGGGGGTT 291
XX |
XX QY 134 GACATTATTTACTGATATGTAATAATAATTTGTATAAGAAATG 179
XX |
XX Db 292 GTTATTATTTACTGATATGTAATAATAATTTGTATAAGAAATG 337
XX |
XX
XX RESULT 15
XX ABL40499
XX ID ABL40499 standard; DNA; 2267 BP.
XX
XX AC ABL40499;
XX
XX DT 17-JUN-2002 (first entry)
XX
XX DE DNA construct B comprising Savinase-CI-2A fusion sequence.
XX
XX KW Protease; allergenicity; detergent; subtilase; S11; S12; additive;
XX protease inhibitor; barley; chymotrypsin; CI-2A; fusion protein;
XX subtilisin; savinase; apr gene; ds.
XX
XX OS Synthetic.
XX OS Bacillus sp.
XX OS Hordeum sp.
XX
XX PN WO200216619-A1.
XX
XX PD 28-FEB-2002.
XX
XX PF 09-JUL-2001; 2001WO-DK000479.
XX
XX PR 21-AUG-2000; 2000DK-00001233.
XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 00:23:20 ; Search time 1078.5 Seconds
(without alignments)
7434.827 Million cell updates/sec

Title: US-09-834-271A-27
Perfect score: 185
Sequence: 1 gsccttaaggcctgaatc.....tgtataagaatggagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sv:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hgtg_hum:*
40: em_hgtg_mus:*
41: em_hgtg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	100.0	185	6	AR160503 Sequence
2	185	100.0	185	6	AX088921 Sequence
3	185	100.0	185	6	BD136799 Process f
4	183.4	99.1	185	6	AR160502 Sequence
5	183.4	99.1	185	6	AX088920 Sequence
6	183.4	99.1	185	6	BD136798 Process f
7	180.2	97.4	185	6	AR160501 Sequence
8	180.2	97.4	185	6	AX088919 Sequence
9	180.2	97.4	185	6	BD136797 Process f
10	157.2	85.0	249	6	BD174625 Modified
11	157.2	85.0	2604	6	AR168304 Sequence
12	157.2	85.0	2604	6	BD249244 Alpha-amy
13	157.2	85.0	2604	6	AR215288 Sequence
14	157.2	85.0	2604	6	AX036892 Sequence
15	157.2	85.0	2604	6	BD137366 Enzymatic
16	154	83.2	162	6	I40596 Sequence 24
17	154	83.2	162	6	I40867 Sequence 24
18	154	83.2	162	6	I40895 Sequence 24
19	154	83.2	162	6	I56837 Sequence 24
20	154	83.2	537	1	BAAMYL
21	154	83.2	576	6	A00155 Nucleotide
22	154	83.2	576	6	A00156 Nucleotide
23	154	83.2	576	6	A00157 Nucleotide
24	154	83.2	576	6	A00158 Nucleotide
25	154	83.2	576	6	A00604 B.amyloliqu
26	154	83.2	576	6	A00605 B.amyloliqu
27	154	83.2	2084	1	BACRAM
28	154	83.2	2084	6	AR08285 Sequence
29	154	83.2	2084	6	AR037275 Sequence
30	154	83.2	2084	6	AR052145 Sequence
31	154	83.2	2084	6	AR087551 Sequence
32	154	83.2	2084	6	AR129916 Sequence
33	154	83.2	2084	6	AR137905 Sequence
34	154	83.2	2084	6	AR143262 Sequence
35	154	83.2	2084	6	BD230131 Alpha-amy
36	154	83.2	2084	6	AR224261 Sequence
37	154	83.2	2084	6	AR285306 Sequence
38	154	83.2	2084	6	AR428861 Sequence
39	154	83.2	2084	6	AX244195 Sequence
40	154	83.2	2084	6	AX305020 Sequence
41	154	83.2	2084	6	AX339268 Sequence
42	154	83.2	2084	6	AX370717 Sequence
43	154	83.2	2084	6	AX601810 Sequence
44	154	83.2	2084	6	BD077082 Alpha-amy
45	154	83.2	2084	6	BD085805 Alpha-amy

ALIGNMENTS

RESULT 1
AR160503
LOCUS
DEFINITION Sequence 27 from patent US 6255076.
ACCESSION AR160503
VERSION AR160503.1 GI:16224480
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner,W., Sloma,A. and Thomas,M.D.
TITLE Methods for producing a polypeptide in a Bacillus cell
JOURNAL Patent: US 6255076-A 27 03-JUL-2001;
FEATURES Location/Qualifiers

AR160503 185 bp DNA linear PAT 17-OCT-2001

```
source 1..185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.8e-36;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCACAGCTGTCGCTGTGTAAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCACAGCTGTCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 2
AX088921 195 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 22 from Patent WO0114534.
ACCESSION AX088921
KEYWORDS AX088921.1 GI:13397680
SOURCE Bacillus subtilis
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
AUTHORS Thomas, M.D. and Brown, K.M.
TITLE Polypeptides having pectin acetylesterase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 22 01-MAR-2001;
FEATURES Novozymes Biotech, Inc. (US)
source Location/Qualifiers
1..185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.8e-36;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCACAGCTGTCGCTGTGTAAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCACAGCTGTCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
BD136799 185 bp DNA linear PAT 18-SEP-2002
LOCUS
```

```
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136799
VERSION BD136799.1 GI:23231744
KEYWORDS JP 2002504379-A/27.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner, W., Sloma, A. and Thomas, M.D.
TITLE Process for producing polypeptide in bacillus cells
JOURNAL Patent: JP 2002504379-A 27 12-FEB-2002;
COMMENT NOVO NORDISK BIOTECH INC
OS Bacillus
PN JP 2002504379-A/27
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/09, C12N1/21, C12N9/54//C07K14/325, (C12N1/21, C12R1:07),
PC C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source 1..185
Location/Qualifiers
1..185
/organism="Bacillus"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.8e-36;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCACAGCTGTCGCTGTGTAAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCACAGCTGTCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
AR160502 185 bp DNA linear PAT 17-OCT-2001
LOCUS Sequence 26 from patent US 6255076.
ACCESSION AR160502
VERSION AR160502.1 GI:16224477
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner, W., Sloma, A. and Thomas, M.D.
TITLE Methods for producing a polypeptide in a Bacillus cell
JOURNAL Patent: US 6255076-A 26 03-JUL-2001;
FEATURES Location/Qualifiers
1..185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
```


Best Local Similarity 99.5%; Pred. No. 2.4e-35; Mismatches 1; Indels 0; Gaps 0;
Matches 184; Conservative 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCATATAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCATATAAATACCTTGT 60

QY 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120

QY 121 ATAAAGGGGGTGTGACATATTTTACTGATATGATATATAATTTGATAGAAATAGG 180
Db 121 ATAAAGGGGGTGTGACATATTTTACTGATATGATATATAATTTGATAGAAATAGG 180

QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 5
AX088920
LOCUS AX088920 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 21 from Patent WO0114534.
ACCESSION AX088920
VERSION AX088920.1 GI:13397679
KEYWORDS
SOURCE Bacillus subtilis
ORGANISM Bacillus subtilis
REFERENCE 1
AUTHORS Thomas,M.D. and Brown,K.M.
TITLE Polypeptides having pectin acetyltransferase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 21 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source Location/Qualifiers
1. .185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.4e-35;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCATATAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCATATAAATACCTTGT 60

QY 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120

QY 121 ATAAAGGGGGTGTGACATATTTTACTGATATGATATATAATTTGATAGAAATAGG 180
Db 121 ATAAAGGGGGTGTGACATATTTTACTGATATGATATATAATTTGATAGAAATAGG 180

QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 6
BD136798
LOCUS BD136798 185 bp DNA linear PAT 18-SEP-2002
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136798
VERSION BD136798.1 GI:23231743
KEYWORDS JP 2002504379-A/26.
SOURCE unidentified
ORGANISM unidentified

unclassified.
1 (bases 1 to 185)
Widner,W., Sloma,A. and Thomas,M.D.
Process for producing polypeptide in bacillus cells
Patent: JP 2002504379-A 26 12-FEB-2002;
NOVO NORDISK BIOTECH INC
OS Bacillus
PN JP 2002504379-A/26
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER,ALAN SLOMA,MICHAEL D THOMAS
PC C12N15/09,C12N1/21,C12N9/54//C07K14/325,(C12N1/21,C12K1:07),
C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source Location/Qualifiers
1. .185
/organism="Bacillus".

FEATURES
source Location/Qualifiers
1. .185
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32844"

ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.4e-35;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCATATAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCATATAAATACCTTGT 60

QY 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120

QY 121 ATAAAGGGGGTGTGACATATTTTACTGATATGATATATAATTTGATAGAAATAGG 180
Db 121 ATAAAGGGGGTGTGACATATTTTACTGATATGATATATAATTTGATAGAAATAGG 180

QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 7
AR160501
LOCUS AR160501 185 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 25 from patent US 6255076.
ACCESSION AR160501
VERSION AR160501.1 GI:16224474
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
1 (bases 1 to 185)
Widner,W., Sloma,A. and Thomas,M.D.
Methods for producing a polypeptide in a Bacillus cell
Patent: US 6255076-A 25 03-JUL-2001;
JOURNAL Location/Qualifiers
source Location/Qualifiers
1. .185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 97.4%; Score 180.2; DB 6; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.5e-34;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCATATAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCATATAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCATTTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185
RESULT 8
AX088919
LOCUS 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 20 from Patent WO0114534.
ACCESSION AX088919
VERSION AX088919.1 GI:13397678
KEYWORDS Bacillus subtilis
SOURCE Bacillus subtilis
ORGANISM Bacillus subtilis
REFERENCE 1
AUTHORS Thomas, M.D. and Brown, K.M.
TITLE Polypeptides having pectin acetylesterase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 20 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source 1.185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"
ORIGIN
Query Match 97.4%; Score 180.2; DB 6; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.5e-34;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGGCCCTGCAATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCCTGCAATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCATTTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185
RESULT 9
BD136797
LOCUS 185 bp DNA linear PAT 18-SEP-2002
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136797
VERSION BD136797.1 GI:23231742
KEYWORDS JP 2002504379-A/25.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner, W., Sloma, A. and Thomas, M.D.
TITLE Process for producing polypeptide in bacillus cells
JOURNAL Patent: JP 2002504379-A 25. 12-FEB-2002;
NOVO NORDISK BIOTECH INC

OS Bacillus
FN JP 2002504379-A/25
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/09, C12N1/21, C12N9/54, C07K14/325, (C12N1/21, C12R1:07),
C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source 1.185
FT Location/Qualifiers
1.185
/organism="Bacillus".
1.185
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 97.4%; Score 180.2; DB 6; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.5e-34;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGGCCCTGCAATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCCTGCAATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCATTTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185
RESULT 10
BD174625
LOCUS 249 bp DNA linear PAT 18-MAR-2003
DEFINITION Modified promoter.
ACCESSION BD174625
VERSION BD174625.1 GI:29120315
KEYWORDS JP 2002272466-A/14.
SOURCE Bacillus amyloliquefaciens
ORGANISM Bacillus amyloliquefaciens
REFERENCE 1 (bases 1 to 249)
AUTHORS Mizubuchi, H., Fushimi, N. and Miyoshi, S.
TITLE Modified promoter
JOURNAL Patent: JP 2002272466-A 14 24-SEP-2002;
SHOWA SANGYO CO LTD
COMMENT OS Bacillus amyloliquefaciens
FN JP 2002272466-A/14
PD 24-SEP-2002
PF 15-MAR-2001 JP 2001074780
PI HIROYUKI MIZUBUCHI, NAOKA FUSHIMI, SHINSUKE MIYOSHI
PC C12N15/09, C12N1/21, C12P21/02, (C12N1/21, C12R1:07), (C12N1/21, PC
C12R1:01),
PC (C12N1/21, C12R1:19), C12N15/00
CC Modified promoter
FH Key Location/Qualifiers
FT source 1.249
FT Location/Qualifiers
1.249
/organism="Bacillus amyloliquefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:1390"
ORIGIN

```
Query Match      85.0%; Score 157.2; DB 6; Length 249;
Best Local Similarity 98.1%; Pred. No. 7.4e-29;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
DB 73 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 132
QY 78 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGACA 137
DB 133 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGTTA 192
QY 138 TTATTTTACTGATGATGATATATATATTTGTTATAGAAAATG 179
DB 193 TTATTTTACTGATGATGATATATATATTTGTTATAGAAAATG 234

RESULT 11
AR168304 AR168304 2604 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 3 from patent US 6287826.
DEFINITION AR168304
ACCESSION AR168304
VERSION AR168304.1 GI:17904148
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..2604
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match      85.0%; Score 157.2; DB 6; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.3e-29;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
DB 593 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 652
QY 78 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGACA 137
DB 653 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGTTA 712
QY 138 TTATTTTACTGATGATGATATATATATTTGTTATAGAAAATG 179
DB 713 TTATTTTACTGATGATGATATATATTTGTTATAGAAAATG 754

RESULT 12
BD249244
LOCUS Alpha-amylase variants.
DEFINITION BD249244
ACCESSION BD249244
VERSION BD249244.1 GI:33059014
KEYWORDS JP 2002540785-A/3.
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Bacillus amyloliquefaciens
PN JP 2002540785-A/3

Query Match      85.0%; Score 157.2; DB 6; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.3e-29;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
DB 593 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 652
QY 78 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGACA 137
DB 653 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGTTA 712
QY 138 TTATTTTACTGATGATGATATATATATTTGTTATAGAAAATG 179
DB 713 TTATTTTACTGATGATGATATATATTTGTTATAGAAAATG 754

RESULT 13
AR215288 AR215288 2604 bp DNA linear PAT 25-SEP-2002
LOCUS Sequence 5 from patent US 6410295.
DEFINITION AR215288
ACCESSION AR215288
VERSION AR215288.1 GI:23313486
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..2604
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match      85.0%; Score 157.2; DB 6; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.3e-29;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
DB 593 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 652
QY 78 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGACA 137
DB 653 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGTTA 712
QY 138 TTATTTTACTGATGATGATATATATTTGTTATAGAAAATG 179
DB 713 TTATTTTACTGATGATGATATATATTTGTTATAGAAAATG 754

RESULT 14
AR215288 AR215288 2604 bp DNA linear PAT 25-SEP-2002
LOCUS Sequence 5 from patent US 6410295.
DEFINITION AR215288
ACCESSION AR215288
VERSION AR215288.1 GI:23313486
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..2604
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match      85.0%; Score 157.2; DB 6; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.3e-29;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
DB 593 ATCGATTGTTGAGAAAAGAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 652
QY 78 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGACA 137
DB 653 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTTGTTA 712
QY 138 TTATTTTACTGATGATGATATATATTTGTTATAGAAAATG 179
DB 713 TTATTTTACTGATGATGATATATATTTGTTATAGAAAATG 754
```

Db 653 ATTTTATGCTGCCAGCTGTCGGCTGTGTAATAAAGAGGAGGTTGTTA 712

Qy 138 TTATTTTACTGATGATATATATATATTTGATAGAAATG 179
Db 713 TTATTTTACTGATGATATATATATATTTGATAGAAATG 754

RESULT 14

AX036892
LOCUS AX036892 2604 bp DNA linear PAT 16-NOV-2000
DEFINITION Sequence 5 from Patent WO0060059.
ACCESSION AX036892
VERSION AX036892.1 GI:11226326
KEYWORDS

ORGANISM

Source
Bacillus amyloliquefaciens
Bacillus amyloliquefaciens
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

REFERENCE
AUTHORS Svendsen, A., Andersen, C., Bisgaard-Frantzen, H., Joergensen, C.T. and Kjaerulff, S.
TITLE Alpha-amylase variants
JOURNAL Patent: WO 0060059-A 5 12-OCT-2000;
NOVONORDISK AS (DK)

FEATURES
source
1..2604
Location/Qualifiers
/organism="Bacillus amyloliquefaciens"
/mol_type="unassigned DNA"
/db_xref="taxon:1390"
707..712
-10_signal
729..734
RBS
759..762
sig_peptide
770..862
CDS
863..2314
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC16484.1"
/db_xref="GI:11226327"
/db_xref="REMBL:CAC16484"
/translation="VNGTLMQYFEWYTPNDQHWKRLONDAEHLSDIGITAVNIPYAY
KGLSQSDNGYGYDILDLGERFQKGTVRKIGTQKSELQDAIGLSHRSNVQYGVDLN
HKAGADATEDVAVEPANRNETSEYQIKAWTDFFRPGNNTSYDFKWHWYHFDG
ADWESRKRISIRFKPRGEGKADWESSENGNYDLYMADVDDYDHPDVAETKKGWIK
YANELSLDGFRIIDAAKHIFKFLRDWQAVRQATKEMFTVAEYQNNAGKLENLYNK
TSFNOSVDFPLHNLQAASSQGGYDMRLLDGTIVSRHPEKAVTENVNHDTPQGS
LESTVQWFKPLAVAFILTRSGYPOVFYGMCTKGTSPKEIPSLKNDTEPILKARK
EYAGPQHDYIDHPDVIGTWREGDSSAAKSLAALIITDGGGSKRWYAGLKNAGETWY
DITGRSDTVIKIGSDGWEHFVNDGVSIVVQK"
863..2311
mat_peptide
/product="unnamed"
2321..2376
terminator

ORIGIN

Query Match 85.0%; Score 157.2; DB 6; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.3e-29;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTCAGAAAAGAGAGACCAATATACTTGTCTGTCTCATCAGACAGGT 77

Db 593 ATCGATTGTTTCAGAAAAGAGAGACCAATATACTTGTCTGTCTCATCAGACAGGT 652

Qy 78 ATTTTATGCTGCCAGACTGCCCTGTGTATAAATAAGGAGGTTGACA 137

Db 653 ATTTTATGCTGCCAGACTGCCCTGTGTATAAATAAGGAGGTTGTTA 712

Qy 138 TTATTTTACTGATGATATATATATTTGATAGAAATG 179

Db 713 TTATTTTACTGATGATATATATTTGATAGAAATG 754

RESULT 15

BD137366

LOCUS

BD137366

DEFINITION

Enzymatic preparation of glucose syrup from starch.

ACCESSION

BD137366

VERSION

JP 2002505885-A/2.

KEYWORDS

Bacillus amyloliquefaciens

ORGANISM

Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

REFERENCE

1 (bases 1 to 2604)

AUTHORS

Norman, B.E. and Hendriksen, H.V.

TITLE

Enzymatic preparation of glucose syrup from starch

JOURNAL

Patent: JP 2002505885-A 2 26-FEB-2002;

NOVOZYMES AS

COMMENT

OS Bacillus amyloliquefaciens

PN JP 2002505885-A/2

PD 26-FEB-2002

PF 08-MAR-1999 JP 2000535766

PR 09-MAR-1998 DK 0321/98

PI BARRIE EDMUND NORMAN HANNE VANG HENDRIKSEN

PC C12P19/14, C12N9/26, C12N15/09, C13K1/06, C12N15/00 CC Enzymatic

preparation of glucose syrup from starch FH Key

Location/Qualifiers

FT -10_signal (707)..(712)

FT -35_signal (729)..(734)

FT RBS (759)..(762)

FT sig_peptide (770)..(862)

FT mat_peptide (863)..(2314)

FT terminator (2321)..(2376)

FT CDS (770)..(2314).

FEATURES

source

1..2604

/organism="Bacillus amyloliquefaciens"

/mol_type="genomic DNA"

/db_xref="taxon:1390"

ORIGIN

Query Match 85.0%; Score 157.2; DB 6; Length 2604;

Best Local Similarity 98.1%; Pred. No. 5.3e-29;

Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTCAGAAAAGAGAGACCAATATACTTGTCTGTCTCATCAGACAGGT 77

Db 593 ATCGATTGTTTCAGAAAAGAGAGACCAATATACTTGTCTGTCTCATCAGACAGGT 652

Qy 78 ATTTTATGCTGCCAGACTGCCCTGTGTATAAATAAGGAGGTTGACA 137

Db 653 ATTTTATGCTGCCAGACTGCCCTGTGTATAAATAAGGAGGTTGTTA 712

Qy 138 TTATTTTACTGATGATATATATATTTGATAGAAATG 179

Db 713 TTATTTTACTGATGATATATATTTGATAGAAATG 754

Search completed: March 19, 2004, 02:14:35

Job time: 1078.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 01:21:15 ; Search time 1837 Seconds
(without alignments)
3007.351 Million cell updates/sec

Title: US-09-834-271A-26
Perfect score: 185
Sequence: 1 ggccttaaggccctcaatc.....tgtataagaagaatggagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pin:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_man:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vtl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	24.3	427	14	CA819917 sau82a05.
C 2	40.6	21.9	1201	9	AL554700 AL554700
C 3	40.4	21.8	735	29	CB808066 tigr-gss-
C 4	40.2	21.7	1201	13	EX446359 EX446359

5	39.2	21.2	699	28	BZ072552
6	38.6	20.9	1038	29	CNS01TL7
7	38.4	20.8	700	28	BH922535
8	38.2	20.6	1201	13	EX356191
9	37.6	20.3	392	10	AW292419
10	37.6	20.3	461	9	AW001963
11	37.6	20.3	462	12	BQ000230
12	37.6	20.3	467	9	AI989386
13	37.6	20.3	475	9	AI697642
14	37.6	20.3	564	9	AI479617
15	37.6	20.3	700	9	AU253100
16	37.2	20.1	868	28	AZ535134
17	37.2	20.1	887	28	AZ546631
18	37.2	20.1	1201	13	EX385059
19	37.2	20.1	1201	13	EX419991
20	37	20.0	509	10	BE157233
21	37	20.0	686	28	BZ028517
22	37	20.0	727	29	CC920927
23	37	20.0	1010	10	BE614304
24	37	20.0	1304	28	CC317939
25	36.8	19.9	523	28	B25666
26	36.8	19.9	1201	9	AL543368
27	36.6	19.8	403	9	AI499797
28	36.6	19.8	876	28	AZ691916
29	36.6	19.8	879	28	AZ672806
30	36.6	19.8	879	28	AZ672806
31	36.6	19.8	1201	9	AL513973
32	36.4	19.7	505	28	AQ460935
33	36.4	19.7	1124	13	EX436382
34	36.4	19.7	1200	13	EX426107
35	36.2	19.6	867	28	BZ446197
36	36.2	19.6	916	14	CA471136
37	36.2	19.6	928	29	CNS000KY
38	36.2	19.6	1101	29	CNS000HP5
39	36	19.5	565	12	BM153720
40	36	19.5	986	29	CNS000EH
41	36	19.5	1040	29	CNS01820
42	36	19.5	1062	28	CC305992
43	35.8	19.4	418	12	BM959532
44	35.8	19.4	453	29	CE516519
45	35.8	19.4	894	9	AL543011

ALIGNMENTS

RESULT 1
CA819917
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
CA819917
sau82a05.y1 Gm-cl048 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl048-2314 5', mRNA sequence.
CA819917
GI:26268854
EST.
Glycine max (soybean)
Glycine max
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 427)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marita,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

CA819917
427 bp mRNA linear EST 09-DEC-2002
sau82a05.y1 Gm-cl048 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl048-2314 5', mRNA sequence.
CA819917
GI:26268854
EST.
Glycine max (soybean)
Glycine max
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 427)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marita,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: cc@resgen.com web site:
 www.resgen.com

Seq primer: -40RP from Gibco
 High quality sequence stop: 418.
 Location/Qualifiers

FEATURES

source

1. 427
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOBEAN CLONE ID: Gm-cl048-2314"
 /tissue_type="whole seedlings of greenhouse grown plants"
 /dev_stage="11 week old"
 /lab_host="DH10B"
 /clone_lib="Gm-cl048"
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The Clark NIL was constructed and seed was provided
 by Dr. J. Specht, University of Nebraska (Shoemaker and
 Specht, 1995). The cDNA library was constructed from mRNA
 isolated from whole seedlings of 1 week old greenhouse
 grown plants. Complementary DNA was synthesized from mRNA
 using a primer consisting of a poly(dT) sequence with a
 XhoI restriction site and a 3' anchor. EcoRI adapters were
 ligated to the blunt-ended cDNA fragments followed by XhoI
 digestion. The cDNA fragments were directionally cloned
 into the EcoRI-XhoI restriction site of the pBluescript
 vector. The ligated cDNA fragments were transformed into
 DH10B host cells (GibcoBRL). The library was constructed
 in cooperation with Dr. Paul Keim's laboratory at Northern
 Arizona University."

ORIGIN

Query Match 24.3%; Score 45; DB 14; Length 427;

Best Local Similarity 54.5%; Pred. No. 0.89; Mismatches 75; Indels 0; Gaps 0;
 Matches 90; Conservative 0

QY 19 TCGATTGTTGAGAAAGAGAACCATATAATCTGTCTCATCAGACAGGTA 78
 Db 164 TTGCCTTTTAAAGGGAAGGGGAAAAAACCTTTTAACTTTTAACTC 223
 QY 79 TTTTATGCTGTCAGACTGTCGCTGTGTAAAGAAAGAAATAGGGGGTTGACAT 138
 Db 224 TTTTCCCAATCCCTTTTTCCTCAAGAGAGAAAAAAGCCCTTTTAT 283
 QY 139 TATTCTACTGATGTATAATATTTGTATAAGAAATGAGC 183
 Db 284 TTTTCTTAAATTTTATTTCAATTTTATAGAAATGGACC 328

RESULT 2
 AL554700/c 1201 bp mRNA linear EST 31-MAY-2003
 LOCUS AL554700 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0DI085YH07 5-PRIME, mRNA sequence.

ACCESSION AL554700
 VERSION AL554700.2 GI:31276510
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL On Feb 15, 2001 this sequence version replaced gi:12895734.
 COMMENT

Contact: Genoscope
 Genoscope, Centre National de Sequencage
 BP 191 91006 EVRY cedex - France

Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 8225.r
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DI085CD04QP1&cluster=8225.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DI085CD04QP1.
 Location/Qualifiers

FEATURES

source

1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DI085YH07"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 21.9%; Score 40.6; DB 9; Length 1201;

Best Local Similarity 39.3%; Pred. No. 6.8; Mismatches 58; Indels 0; Gaps 0;
 Matches 59; Conservative 33

QY 17 AATCGATTCTTTGAGAAAGAGAACCATATAATCTGTCTCATCAGACAGG 76
 Db 1010 AAAAATTTTTTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 951
 QY 77 TATTTTATGCTGTCAGACTGTCGCTGTGTAAAGAAAGAAATAGGGGGTTGAC 136
 Db 950 WAWTTTTTTTTTTTTTTTTTTTTTCTVVRVRRRRRAAARDDGGGKKKT 891
 QY 137 ATATTTTACTGATGTATAATATAATT 166
 Db 890 NTTTTTTTTTTTTTTAAATGATACACTT 861

RESULT 3
 CE808066
 LOCUS tigr-gss-dog-17000317853400 Dog Library Canis familiaris genomic,
 DEFINITION genomic survey sequence.

ACCESSION CE808066
 VERSION CE808066.1 GI:37149050
 KEYWORDS GSS.
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 1 (bases 1 to 735)
 Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.
 The dog genome: survey sequencing and comparative analysis
 Science 301 (5641), 1898-1903 (2003)

TITLE The dog genome: survey sequencing and comparative analysis
 JOURNAL Science
 MEDLINE 22875432
 PUBMED 14512627
 COMMENT Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirkness@tigr.org
 Class: shotgun.
 Location/Qualifiers

1. 735
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"

FEATURES

source

```

/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match          21.8%; Score 40.4; DB 29; Length 735;
Best Local Similarity 53.9%; Pred. No. 8.9;
Matches 83; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 30 AGAAGAAGAGAGACATATAATACCTTGCTGTCATCAGACAGGATTTTATGCT 89
Db 15 ACAAAAACAAACAAACAAATTCCTGTTTTCCTCTTTTCTTTTCTTTTCTTTT 74
QY 90 GTCCAGACTGCGCTGTGTAAAGAGGATTAAGGGGGTTCACATTTTACTGA 149
Db 75 TTTTGTGTTTCTCTCTGTTGTAATAAATGAATGAACACCACTTTGTTTGTGTTT 134
QY 150 TATGATATAATAATTTGATAGAAATGGAGC 183
Db 135 GTTATTATTATTATAGACCAAGAGAGGAGC 168

RESULT 4
BX446359/c
LOCUS
DEFINITION
BX446359 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA003ZG04
3-PRIME mRNA sequence.
ACCESSION
BX446359
VERSION
BX446359.1 GI:31023719
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 1201)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2980.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBA003ZG04P1&cluster=2980.r. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA003ZG04P1.

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA003ZG04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and ScaRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match          21.7%; Score 40.2; DB 13; Length 1201;
Best Local Similarity 40.5%; Pred. No. 8.4;
Matches 70; Conservative 27; Mismatches 76; Indels 0; Gaps 0;

QY 5 TTAAGCGCTGCAATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTCTGCT 64
Db 1024 TGAARAWTTCGGTGGGTGGGGGAGAAACCCSYGTCGTGAACMCCMCACWAA 965

```

```

65 CATCAGACAGGGTATTTTATCTGTCAGACTGTCGCTGTGTATATAAAAGGAATAA 124
Db 964 CACGAAACAGATGKTTGTGCGCTGTGCGGWTGTTTTTCTGKGRADAAAARAWW 905
QY 125 AGGGGGTTCACATTTTACTGATATGATATAATAATTTGTATAAGAAA 177
Db 904 CTGGGGCTGYTYYTGTGAPAAAAAAAATAAAWAACCTCCANAWAA 852

RESULT 5
BZ072552
LOCUS
DEFINITION
BZ072552 699 bp DNA linear GSS 10-OCT-2002
sequence.
ACCESSION
BZ072552
VERSION
BZ072552.1 GI:23690695
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 699)
Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,
Nash, W., Rabinowicz, P.D. and Wilson, R.K.
Whole genome shotgun reads from Brassica oleracea
Unpublished (2002)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: lkf46 row: b column: 09
Seq primer: -21UPPOT forward
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 551.
Location/Qualifiers
1..699
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B. oleracea002"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN
Query Match          21.2%; Score 39.2; DB 28; Length 699;
Best Local Similarity 53.2%; Pred. No. 17;
Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 22 ATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGTCTGTCATCAGACGGTATTT 81
Db 328 ATTGTTAAAAATTATATAATAAAATAAAATAATTTAATTGTAATTTATGTTATTT 387
QY 82 TTTATGCTGCCAGACTGTCGCTGTGTATAAAAGGAATAAAGGGGGTTCACATTAT 141
Db 388 TTAATATTGTGAAGACTAGATTTTGTAGAAACTGAGGAAATACCATGTTATGTCAT 447
QY 142 TTTACTGATATGATATAATAATTTGTATAAGAAA 177
Db 448 TTTATAAATGAATTAAGATTATAGATTATTACAA 483

RESULT 6
CNS01TL7
LOCUS
DEFINITION
CNS01TL7 1038 bp DNA linear GSS 01-SEP-2000
Tetradon nigroviridis genome survey sequence T7 end of clone

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

ORIGIN

RESULT 8
EX356191
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

GSS.
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 700)
 Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
 Nash,W., Rabinowicz,P.D. and Wilson,R.K.
 Whole genome shotgun reads from Brassica oleracea
 Unpublished (2002)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Plate: odi94 row: C column: 08
 Seq primer: -21UPPOT forward
 Class: shotgun
 High quality sequence start: 56
 High quality sequence stop: 551.
 Location/Qualifiers
 1..700
 /organism="Brassica oleracea"
 /mol_type="Genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea002"
 /note="vector: pOTW13; Whole genome shotgun library from
 flowering buds. DNA was purified from a crude nuclear
 prep using Brassica oleracea TO1000DP3 buds provided by
 Thomas Osborn at the University of Wisconsin. Genomic
 DNA was provided by Pablo Rabinowicz (CSHL) and the
 shotgun library prepared at Washington University Genome
 Sequencing Center."
 Query Match 20.8%; Score 38.4; DB 28; Length 700;
 Best Local Similarity 53.3%; Pred. No. 27;
 Matches 81; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
 QY 26 TTTGAGAAAAGAGAAGACCATAAAAAATACCTTGTCTGTCATCAGACAGCGGTATTTTTTA 85
 Db 313 TTTTACAAAAAAATTTTGTATATAATTCGATTTTTCGATTTTACAAAACATTTTTTA 372
 QY 86 TCGCTGCCAGCTGCCGTGTGTAATAAAGGAATAAGGGGGTTCACATTATTTTA 145
 Db 373 ATATCTATACACTTTTTTTTGTATAAACATATATTTCGGATTTTAAATATTTT 432
 QY 146 CTGATATGATATAATAATTGTATAGAAA 177
 Db 433 AATATATATATATATATATATATATATTATAA 464
 BX356191 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 clone CS0D1009YC24 5-PRIME, mRNA sequence.
 BX356191
 BX356191.1 GI:30365992
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 31006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 1433.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI009BB120P1&cluster=1433.r>. Contact : Peng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> InvitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DI009BB120P1. Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI009YC24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dr) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

ORIGIN

Query Match 20.6%; Score 38.2; DB 13; Length 1201;
Best Local Similarity 45.2%; Pred. No. 25;
Matches 70; Conservative 16; Mismatches 69; Indels 0; Gaps 0;

Qy 23 TTGTTGAGAAAGAGACACATAAATACCTGTCTGTCATCAGACAGGGTATTTT 82
Db 892 TTTTAAAAAARAAAAAATAAATAATATATATTTTGGRAWTTTRAGWTTT 951

Qy 83 TTATGCTGCCAGCTGTCGCTGTGTATAAAAAAGGAATAAGCGGGTGTGACATTTT 142
Db 952 TTTTATATATATAGATTTTAAAAAATAAATAAAGCGGCGGAGTWTWTWTW 1011

Qy 143 TTACTCATATGATAATAATTTGTATAGAAAA 177
Db 1012 TTTTTTTTTTTTTTTTTTTTTTTTTWWAAAAAAA 1046

RESULT 9

AW292419 392 bp mRNA linear EST 16-JAN-2000
LOCUS UI-H-B12-agw-h-08-0-UI.s1 NCI_CGAP_Sub4 Homo sapiens CDNA clone
DEFINITION IMAGE:2725886 3', mRNA sequence.

AW292419
AW292419.1 GI:6699055
VERSION
ACCESSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html The following repetitive elements were found in this cDNA sequence: 1-30,
>AT rich/Low complexity
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source

1. 392
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:2725886"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_Sub4"
/note="Vector: pT773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NCI_CGAP_Sub4 library is a subcloned library derived from the NCI_CGAP_Sub2 library which is a subcloned library derived from the NCI_CGAP_Sub1 library, which is a subcloned library derived from Bi. Bi constitutes a mixture of 21 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_C04, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_C010, NCI_CGAP_C016, NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Ly42, NCI_CGAP_Br2, NCI_CGAP_C08, NCI_CGAP_CLL1, NCI_CGAP_Ly42, NCI_CGAP_Br23, NCI_CGAP_Ly5, NCI_CGAP_Ly44, NCI_CGAP_Ly9, NCI_CGAP_GC4, NCI_CGAP_GC6, NCI_CGAP_Br25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1 : LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE cloneIDs 132376-132391),
1456008-1456775, 1500552-1502853) NCI_CGAP_Kid5 pool 1 :
LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE cloneIDs
1323912-1325831, 1471368-1472903, 1492104-1493255)
NCI_CGAP_Ly5 pool 1 : LLAM 3575-3582, 3851-3854 (IMAGE
cloneIDs 1414920-1417991, 1520904-1522439) NCI_CGAP_GC4
pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE
cloneIDs 1257096-1258631, 1459064-1470983,
1475592-1476743) NCI_CGAP_Pr22 pool 1 : LLAM 2457-2459,
2758-2759, 3062-3068 (IMAGE cloneIDs 985608-986759,
1101192-1101959, 1217928-1220615) NCI_CGAP_C010 pool 1 :
LLAM 2644-2653, 2871-2872 (IMAGE cloneIDs 1057416-1061255,
114584-1145351) Subtraction was performed as previously
described [Bonaldo, Lennon & Soares (1996): Normalization
and Subtraction: Two Approaches To Facilitate Gene
Discovery. Genome Research 6, 791-806.]
TAG TISSUE=lung
TAG LIB=NCI_CGAP_Ly5
TAG_SEQ=CAAC"

ORIGIN

Query Match 20.3%; Score 37.6; DB 10; Length 392;
Best Local Similarity 72.1%; Pred. No. 50;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 111 AAAAAAGGAATTAAGGGGGTTCACATTTTCTGATATATATAATTTGTAT 170
Db 193 AAAGAGAGGAGGAACTAAGTTGTCTATTAAATGACAGATAAGTAAAAAATTTGCTTT 252

Qy 171 AAGAAAT 178
Db 253 AAGAAAT 260

RESULT 10

AW001963 461 bp mRNA linear EST 09-MAR-2000
LOCUS WT82301.x1 Soares_thymus_NHTh Homo sapiens CDNA clone
DEFINITION IMAGE:2513953 3', mRNA sequence.

AW001963
AW001963.1 GI:5848879
VERSION
ACCESSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 461)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 696 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 448.

FEATURES

source

1. .461
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2513953"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares_thymus_NHFT"
/note="Organ: thymus, pooled; Vector: p773D-Pac
(Pharmacia) with a modified polylinker; Site: 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo (dT) primer [5',
TGTTACCAATCTGAAGTGGGCGCGCAACGTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 20.3%; Score 37.6; DB 9; Length 461;
Best Local Similarity 72.1%; Pred. No. 48;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAAAGGAATAAGGGGGTGGACATTTTACTGATATGATATATAATTTGTAT 170
|||||
Db 176 AAGAGAGGAGGAACTAAGGTGTCTTAATGACAGATAAGTAAATAATTTGCTT 235
|||||
QY 171 AAGAAAT 178
|||||
Db 236 AAGAAAT 243
|||||

RESULT 11

BQ000230/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

UNPUBLISHED (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

The following repetitive elements were found in this cDNA

sequence: 433-454, >AT-rich#Low_complexity (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

1. .462
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:582595"
/tissue_type="Lung Focal Fibrosis"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP D10"
/note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a
modified polylinker; Site 1: Eco RI; Site 2: Not I;
NCI CGAP D10 is a cDNA library containing the following
tissue(s): A pool of Lung focal Fibrosis. The library was
constructed according to Bonaldo, Lemmon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Eco RI
adaptor, digested with Not I, and cloned directionally
into p773-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ATACGGCGTC.
TAG TISSUE=lung with fibrosis
TAG LIB=UI-H-D10
TAG_SEQ=ATACGGCGTC"

ORIGIN

Query Match 20.3%; Score 37.6; DB 12; Length 462;
Best Local Similarity 72.1%; Pred. No. 48;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAAAGGAATAAGGGGGTGGACATTTTACTGATATGATATATAATTTGTAT 170
|||||
Db 270 AAGAGAGGAGGAACTAAGGTGTCTTAATGACAGATAAGTAAATAATTTGCTT 211
|||||
QY 171 AAGAAAT 178
|||||
Db 210 AAGAAAT 203
|||||

RESULT 12

AI989386

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

UNPUBLISHED (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 442.

Location/Qualifiers

1. .467

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2513806"

/dev_stage="fetal"

/lab_host="DH10B (phage-resistant)"

/clone_lib="Soares_thymus_NHFT"

/note="Organ: thymus, pooled; Vector: p773D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

AI989386 467 bp mRNA linear EST 08-SEP-1999
wt8e12.x1 Soares_thymus_NHFT Homo sapiens cDNA clone
IMAGE:2513806 3', mRNA sequence.

AI989386

AI989386.1 GI:5836309

EST.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 467)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 442.

Location/Qualifiers

1. .467

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2513806"

/dev_stage="fetal"

/lab_host="DH10B (phage-resistant)"

/clone_lib="Soares_thymus_NHFT"

/note="Organ: thymus, pooled; Vector: p773D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5',
TCTTACCATCTGAGTGGAGCGCGCCAGCGTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 20.3%; Score 37.6; DB 9; Length 467;
Best Local Similarity 72.1%; Pred. No. 48;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAAAGGAATAAAGGGGGTTCACATTATTTTACTGATATGATATGATATATATATATTTGTAT 170
|||||
Db 176 AAGAGAGGAGGAACTAGGTGTCATTAAATGACAGATAAGTAAAAAATATTGCTTT 235
|||||

QY 171 AAGAAAT 178
|||||
Db 236 AAGAAAT 243
|||||

RESULT 13
AI697642 475 bp mRNA linear EST 18-DEC-1999
LOCUS wel5902.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2341202 3',
DEFINITION mRNA sequence.

ACCESSION AI697642

VERSION AI697642

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 475)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

EMAIL: cgapbs@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 701 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 465.

Location/Qualifiers

1. .475

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2341202"

/tissue_type="carcinoid"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Lu24"

/notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; plasmid DNA from the normalized

library NCI CGAP Lu5 was prepared, and ss circles were

made in vitro. Following HAP purification, this DNA was

used as tracer in a subtractive hybridization reaction.

The driver was PCR-amplified cDNAs from a pool of 5,000

clones made from the same library (clones

1414920-1417991 and 1520904-152439). Subtraction by Bento

Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 20.3%; Score 37.6; DB 9; Length 475;

Best Local Similarity 72.1%; Pred. No. 47;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAAAGGAATAAAGGGGGTTCACATTATTTTACTGATATGATATATATATATTTGTAT 170
|||||
Db 179 AAGAGAGGAGGAACTAGGTGTCATTAAATGACAGATAAGTAAAAAATATTGCTTT 238
|||||

QY 171 AAGAAAT 178
|||||

Db 239 AAGAAAT 246
|||||

RESULT 14
AI479617 564 bp mRNA linear EST 14-APR-1999
LOCUS tm32812.x1 NCI_CGAP_CELL1 Homo sapiens cDNA clone IMAGE:2158271 3',
DEFINITION mRNA sequence.

ACCESSION AI479617

VERSION AI479617.1 GI:4372785

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 564)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

EMAIL: cgapbs@mail.nih.gov

Tissue Procurement: Ash. Alizadeh, John Byrd, M.D., Mike Grever,

M.D., Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 685 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 446.

Location/Qualifiers

1. .564

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2158271"

/tissue_type="B-cell, chronic lymphocytic leukemia"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_CELL1"

/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5,

TGTTACCATCTGAGTGGAGCGCGCATGCTTTTCTTTTCTTTTCTTTTCTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7T3 vector.

Library is normalized, and was constructed by Bento

Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 20.3%; Score 37.6; DB 9; Length 564;

Best Local Similarity 72.1%; Pred. No. 45;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAAAGGAATAAAGGGGGTTCACATTATTTTACTGATATGATATATATATTTGTAT 170
|||||
Db 173 AAGAGAGGAGGAACTAGGTGTCATTAAATGACAGATAAGTAAAAAATATTGCTTT 232
|||||

QY 171 AAGAAAT 178
|||||

Db 233 AAGAAAT 240
|||||

```

RESULT 15
LOCUS      AU253100
DEFINITION AU253100 human favorable neuroblastoma cDNA Homo sapiens EST 27-AUG-2003
            Nb1a03504 3', mRNA sequence.
ACCESSION  AU253100
VERSION    AU253100.1 GI:34321813
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 700)
AUTHORS   Chirana, M., Morohashi, A., Inuzuka, H., Shishikura, T., Kawamoto, T.,
            Kageyama, H., Nakamura, Y., Isogai, E., Takayasu, K., Sakiyama, S.,
            Suzuki, Y., Sugano, S., Goto, T., Sato, S. and Nakagawara, A.
TITLE     Expression profiling and characterization of 4200 genes cloned from
            primary neuroblastomas: identification of 305 genes differentially
            expressed between favorable and unfavorable subsets
JOURNAL    Oncogene 22 (35), 5525-5536 (2003)
MEDLINE    22815148
PUBMED     12934113
COMMENT    Contact: Akira Nakagawara
            Division of Biochemistry
            Chiba Cancer Research Institute
            666-2 Nitona, Chuo-Ku, Chiba 260-8717, Japan
            Tel: 81-43-264-5431 (ex.5201)
            Fax: 81-43-265-4459
            Email: akirana@chiba-ceri.chuo.chiba.jp.
FEATURES   source
            1..700
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="Nb1a03504"
            /tissue_type="neuroblastoma"
            /clone_lib="human favorable neuroblastoma cDNA"
ORIGIN
Query Match      20.3%; Score 37.6; DB 9; Length 700;
Best Local Similarity 72.1%; Pred. No. 41;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 111 AAAAAAGGAATAAGCGGGGTGACATTATTTTACTGATATGATATATATATATTTGTAT 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163 AAAGAGAGAGGAGAACTAAGGTTGTCATTAAATGACAGATAAGTAAATAATTTGCTTT 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 AAGAAAAAT 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 223 AAGAAAAAT 230

```

Search completed: March 19, 2004, 03:16:12
 Job time : 1842 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 02:14:41 ; Search time 745.5 Seconds
(without alignments)
917.966 Million cell updates/sec

Title: US-09-834-271A-26
Perfect score: 185
Sequence: 1 ggccttaaggccgtgaatc.....tgtataagaatggagctc 185

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:**
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	185	100.0	185	10	US-09-834-271A-26
2	185	100.0	185	14	US-10-406-025-5
3	183.4	99.1	185	10	US-09-834-271A-27
4	183.4	99.1	185	14	US-10-406-025-6
5	178.6	96.5	185	10	US-09-834-271A-25
6	178.6	96.5	185	14	US-10-406-025-7
7	159.6	86.3	2166	12	US-10-344-231-17
8	159.6	86.3	2166	12	US-10-363-332A-17
9	159.6	86.3	2267	12	US-10-344-231-18
10	159.6	86.3	2267	12	US-10-363-332A-18
11	159.6	86.3	2588	12	US-10-344-231-20
12	159.6	86.3	2588	12	US-10-363-332A-20
13	155.6	84.1	2084	9	US-09-769-864-58
14	155.6	84.1	2084	9	US-09-854-346-9
15	155.6	84.1	2084	9	US-09-918-543-9

16	155.6	84.1	2084	10	US-09-925-576C-9	Sequence 9, Appli
17	155.6	84.1	2084	12	US-10-665-667-58	Sequence 58, Appli
18	155.6	84.1	2084	12	US-10-644-187-3	Sequence 3, Appli
19	155.6	84.1	2084	14	US-10-184-771-3	Sequence 3, Appli
20	155.6	84.1	2084	14	US-10-186-042-3	Sequence 3, Appli
21	155.6	84.1	2604	14	US-10-146-327-5	Sequence 5, Appli
22	95.2	51.5	13222	10	US-09-738-116-60	Sequence 60, Appli
23	88.6	47.5	6837	10	US-09-928-847B-49	Sequence 49, Appli
24	44	23.8	44	10	US-09-834-271A-19	Sequence 19, Appli
25	39	21.1	42	10	US-09-834-271A-9	Sequence 10, Appli
26	38.4	20.8	43	10	US-09-834-271A-10	Sequence 10, Appli
C 27	37.6	20.3	2141	15	US-10-104-047-1737	Sequence 1737, Ap
C 28	37	20.0	3673778	14	US-10-312-841-2	Sequence 2, Appli
C 29	36	19.5	28588	9	US-09-764-887-399	Sequence 399, App
C 30	36	19.5	28588	14	US-10-073-961-339	Sequence 399, App
C 31	35.6	19.2	179	9	US-09-783-590-7881	Sequence 7881, Ap
32	35.2	19.0	317	15	US-10-125-968-1260	Sequence 1260, Ap
33	34.8	18.8	3673778	14	US-10-312-841-1	Sequence 1, Appli
34	34.6	18.7	14147	12	US-10-221-744A-470	Sequence 470, App
35	34.6	18.7	14147	14	US-10-172-086-52	Sequence 52, Appl
C 36	34.4	18.6	383	9	US-09-969-347-299	Sequence 299, App
C 37	34.2	18.5	9121	9	US-09-070-927A-221	Sequence 221, App
C 38	34.2	18.5	3673778	14	US-10-312-841-1	Sequence 1, Appli
C 39	33.8	18.3	347	12	US-10-424-599-69129	Sequence 69129, A
C 40	33.8	18.3	474	10	US-09-764-891-1477	Sequence 1477, Ap
41	33.8	18.3	158001	16	US-10-211-179-11	GENERAL INFORMATI
42	33.6	18.2	512	15	US-10-027-632-15736	Sequence 15736, A
43	33.6	18.2	827	15	US-10-027-632-145937	Sequence 145937,
44	33.6	18.2	19380	12	US-10-221-613-399	Sequence 389, App
C 45	33.6	18.2	185695	14	US-10-020-141-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-834-271A-26
; Sequence 26, Application US/09834271A
; Publication No. US20030170876A1
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/834,271A
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US/09/258,377
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
; BEST LOCAL SIMILARITY: 100.0%; Pred. No. 1.3e-41;
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
; US-09-834-271A-26

Query Match	100.0%;	Score 185;	DB 10;	Length 185;
Best Local Similarity	100.0%;	Pred. No. 1.3e-41;		
Matches 185;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGCTTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGACCACTATAAATACCTTGT	60	
Db	1	GGCTTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGACCACTATAAATACCTTGT	60	
Qy	61	CTGTCTATCAGACAGGGTATTTTTTATGCTGTCTCCAGACTGTCGCTGTGTAAGGGA	120	
Db	61	CTGTCTATCAGACAGGGTATTTTTTATGCTGTCTCCAGACTGTCGCTGTGTAAGGGA	120	
Qy	121	ATTAAGGGGGGTGACATTTATTTACTGATATGATATATATATTTGTATAAGAAATGG	180	

Db 121 ATAAAGGGGGTGCATTTTACTGATATGATATATATTTGTTATAGAAATGG 180
181 AGCTC 185
181 AGCTC 185

RESULT 2
US-10-406-025-5
; Sequence 5, Application US/10406025
; Publication No. US20030186380A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030186380Alcozymes Biotech, Inc.
; APPLICANT: Thomas, Michael D.
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparaginase
; FILE REFERENCE: 10289.200-US
; CURRENT APPLICATION NUMBER: US/10/406,025
; PRIOR FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,192
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-10-406-025-5

Query Match 100.0%; Score 185; DB 14; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.3e-41;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTTACTGATATGATATATATTTGTTATAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTTTACTGATATGATATATATTTGTTATAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
US-09-834-271A-27
; Sequence 27, Application US/09834271A
; Publication No. US20030170876A1
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/834,271A
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US/09/258,377
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 185

; TYPE: DNA
; ORGANISM: Bacillus
US-09-834-271A-27

Query Match 99.1%; Score 183.4; DB 10; Length 185;
Best Local Similarity 99.5%; Pred. No. 3.5e-41;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTTACTGATATGATATATATTTGTTATAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTTTACTGATATGATATATATTTGTTATAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
US-10-406-025-6
; Sequence 6, Application US/10406025
; Publication No. US20030186380A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030186380Alcozymes Biotech, Inc.
; APPLICANT: Thomas, Michael D.
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparagina
; FILE REFERENCE: 10289.200-US
; CURRENT APPLICATION NUMBER: US/10/406,025
; PRIOR FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,192
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-10-406-025-6

Query Match 99.1%; Score 183.4; DB 14; Length 185;
Best Local Similarity 99.5%; Pred. No. 3.5e-41;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTTACTGATATGATATATATTTGTTATAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTTTACTGATATGATATATATTTGTTATAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
US-09-834-271A-25
; Sequence 25, Application US/09834271A
; Publication No. US20030170876A1

GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/834,271A
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US/09/258,377
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-834-271A-25

Query Match 96.5%; Score 178.6; DB 10; Length 185;
Best Local Similarity 97.8%; Pred. No. 7.6e-40;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTCGAATCGATTGTTGAGAAAAGAGAGACCATAAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTCGAATCGATTGTTGAGAAAAGAGAGACCATAAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAGGA 120

QY 121 ATAAAGGGGGTGGACATTTTACTGATATGTAATAATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGGTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 6
US-10-406-025-7
; Sequence 7, Application US/10406025
; Publication No. US20030186380A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030186380A1ozymes Biotech, Inc.
; APPLICANT: Thomas, Michael D.
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparaginase
; TITLE OF INVENTION: activity
; FILE REFERENCE: 10289.200-US
; CURRENT APPLICATION NUMBER: US/10/406,025
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,192
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-10-406-025-7

Query Match 96.5%; Score 178.6; DB 14; Length 185;
Best Local Similarity 97.8%; Pred. No. 7.6e-40;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTCGAATCGATTGTTGAGAAAAGAGAGACCATAAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTCGAATCGATTGTTGAGAAAAGAGAGACCATAAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAGGA 120

QY 121 ATAAAGGGGGTGGACATTTTACTGATATGTAATAATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGGTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 7
US-10-344-231-17
; Sequence 17, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 2166
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-344-231-17

Query Match 86.3%; Score 159.6; DB 12; Length 2166;
Best Local Similarity 97.6%; Pred. No. 3.5e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGTCTGTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGTCTGTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAGGAATAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAGGAATAGGGGGTT 291

QY 134 GACATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 179
DB 292 GTTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 337

RESULT 8
US-10-363-332A-17
; Sequence 17, Application US/10363332A
; Publication No. US20040038375A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: No. US20040038375A1regard-Madsen, Mads
; TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
; FILE REFERENCE: 10080.204-US
; CURRENT APPLICATION NUMBER: US/10/363,332A
; CURRENT FILING DATE: 2003-05-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 2166
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Construct A
US-10-363-332A-17

Query Match 86.3%; Score 159.6; DB 12; Length 2166;

Best Local Similarity 97.6%; Pred. No. 3.5e-34; DB 12; Length 2267;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTATAAAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTATAAAAGGAATAAAGGGGGTT 291

QY 134 GACATTTTACTGATATGTAATAATTAATTTGTATAAGAAATG 179
DB 292 GTTATTATTTACTGATATGTAATAATTAATTTGTATAAGAAATG 337

RESULT 9

US-10-344-231-18
; Sequence 18, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 18
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-344-231-18

Query Match 86.3%; Score 159.6; DB 12; Length 2267;
Best Local Similarity 97.6%; Pred. No. 3.6e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTATAAAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTATAAAAGGAATAAAGGGGGTT 291

QY 134 GACATTTTACTGATATGTAATAATTAATTTGTATAAGAAATG 179
DB 292 GTTATTATTTACTGATATGTAATAATTAATTTGTATAAGAAATG 337

RESULT 10

US-10-363-332A-18
; Sequence 18, Application US/10363332A
; Publication No. US20040038375A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
; FILE REFERENCE: 10080.204-US
; CURRENT APPLICATION NUMBER: US/10/363,332A
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 18
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Construct B

US-10-363-332A-18

Query Match 86.3%; Score 159.6; DB 12; Length 2267;
Best Local Similarity 97.6%; Pred. No. 3.6e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTATAAAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTATAAAAGGAATAAAGGGGGTT 291

QY 134 GACATTTTACTGATATGTAATAATTAATTTGTATAAGAAATG 179
DB 292 GTTATTATTTACTGATATGTAATAATTAATTTGTATAAGAAATG 337

RESULT 11

US-10-344-231-20
; Sequence 20, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 20
; LENGTH: 2588
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc feature
; LOCATION: (2070)..(2070)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2077)..(2077)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2084)..(2084)
; OTHER INFORMATION: n is a, c, g, or t
US-10-344-231-20

Query Match 86.3%; Score 159.6; DB 12; Length 2588;
Best Local Similarity 97.6%; Pred. No. 3.8e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTATAAAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTATAAAAGGAATAAAGGGGGTT 291

QY 134 GACATTTTACTGATATGTAATAATTAATTTGTATAAGAAATG 179
DB 292 GTTATTATTTACTGATATGTAATAATTAATTTGTATAAGAAATG 337

RESULT 12

US-10-363-332A-20
; Sequence 20, Application US/10363332A
; Publication No. US20040038375A1

GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: No. US2004003875Alreggaard-Madsen, Mads
; TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
; FILE REFERENCE: 10080.204-US
; CURRENT APPLICATION NUMBER: US/10/363,332A
; CURRENT FILING DATE: 2003-05-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 20
; LENGTH: 2588
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Construct D
; NAME/KEY: misc feature
; LOCATION: (2070)..(2070)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2077)..(2077)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2084)..(2084)
; OTHER INFORMATION: n is a, c, g, or t
; OTHER INFORMATION: n is a, c, g, or t
US-10-363-332A-20

Query Match 86.3%; Score 159.6; DB 12; Length 2588;
Best Local Similarity 97.6%; Pred. No. 3.8e-34; Indels 0; Gaps 0;
Matches 162; Conservative 0; Mismatches 4;
QY 14 TCGAATCGATTGTTGAGAAAGAGAACCCATATAAAATACCTTCTGTCTATCAGACA 73
Db 172 TCGAATCGATTGTTGAGAAAGAGAACCCATATAAAATACCTTCTGTCTATCAGACA 231
QY 74 GGGTATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAGGAAGAAAGAAAGGGGGT 133
Db 232 GGGTATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAGGAAGAAAGAAAGGGGGT 291
QY 134 GACATTATTTTACTGATATGATATATAATATATATATTTGTTATAGAAAATG 179
Db 292 GTTATTTTACTGATATGATATATAATATATATATTTGTTATAGAAAATG 337

RESULT 13
US-09-769-864-58
; Sequence 58, Application US/09769864
; Patent No. US20010039253A1
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kiaerulff, Soren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/769,864
; CURRENT FILING DATE: 2001-01-25
; PRIOR FILING DATE: 09/183,412
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)....(1794)
US-09-769-864-58

Query Match 84.1%; Score 155.6; DB 9; Length 2084;
Best Local Similarity 97.5%; Pred. No. 4.5e-33; Indels 0; Gaps 0;
Matches 158; Conservative 0; Mismatches 4;
QY 18 ATCGATTGTTTGAGAAAGAGAACCCATATAAAATACCTTCTGTCTATCAGACAGGT 77
Db 73 ATCGATTGTTTGAGAAAGAGAACCCATATAAAATACCTTCTGTCTATCAGACAGGT 132
QY 78 ATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAGGAAGAAAGAAAGGGGGTTCACA 137
Db 133 ATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAGGAAGAAAGAAAGGGGGTTCACA 192
QY 138 TTATTTACTGATATGATATATAATATATAATTTGTTATAGAAAATG 179
Db 193 TTATTTACTGATATGATATATAATATATAATTTGTTATAGAAAATG 234

RESULT 14
US-09-854-346-9
; Sequence 9, Application US/09854346
; Patent No. US20020068352A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020068352A1ozymes A/S
; APPLICANT: Svendsen, Allan
; APPLICANT: Jorgensen, Christel Thea
; APPLICANT: Nielsen, Bjarne Romfeldt
; TITLE OF INVENTION: Alpha-amyase variants with altered 1.6 activity
; FILE REFERENCE: 6140.200-US
; CURRENT APPLICATION NUMBER: US/09/854,346
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 9
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)....(1794)
; OTHER INFORMATION: BAN
US-09-854-346-9

Query Match 84.1%; Score 155.6; DB 9; Length 2084;
Best Local Similarity 97.5%; Pred. No. 4.5e-33; Indels 0; Gaps 0;
Matches 158; Conservative 0; Mismatches 4;
QY 18 ATCGATTGTTTGAGAAAGAGAACCCATATAAAATACCTTCTGTCTATCAGACAGGT 77
Db 73 ATCGATTGTTTGAGAAAGAGAACCCATATAAAATACCTTCTGTCTATCAGACAGGT 132
QY 78 ATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAGGAAGAAAGAAAGGGGGTTCACA 137
Db 133 ATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAGGAAGAAAGAAAGGGGGTTCACA 192
QY 138 TTATTTACTGATATGATATATAATATATAATTTGTTATAGAAAATG 179
Db 193 TTATTTACTGATATGATATATAATATATAATTTGTTATAGAAAATG 234

RESULT 15
US-09-918-543-9
; Sequence 9, Application US/09918543
; Patent No. US20020155574A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020155574A1ozymes A/S
; APPLICANT: Thisted, Thomas
; APPLICANT: Kiaerulff, Soren
; APPLICANT: Andersen, Carsten
; APPLICANT: Fuglsang, Claus Crone
; TITLE OF INVENTION: Alpha-amyase mutants with altered properties
; FILE REFERENCE: 10062.200-US
; CURRENT APPLICATION NUMBER: US/09/918,543

Mon Mar 22 09:21:45 2004

```
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)..(1794)
; OTHER INFORMATION:
US-09-918-543-9

Query Match      84.1%; Score 155.6; DB 9; Length 2084;
Best Local Similarity 97.5%; Pred. No. 4.5e-33;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      18 ATCGATTGTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACAGG 77
Db      73 ATCGATTGTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACAGG 132

QY      78 ATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAAGGAATAAAGGGGGTTGACA 137
Db      133 ATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAATAGGATTAAGGGGGTTGCTTA 192

QY      138 TTATTTTACTGATATGTATAATAATTTGTATAAGAAATG 179
Db      193 TTATTTTACTGATATGTATAATAATTTGTATAAGAAATG 234

Search completed: March 19, 2004, 05:16:19
Job time : 753.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 01:23:55 ; Search time 44 Seconds
(without alignments)
2333.316 Million cell updates/sec

Title: US-09-834-271A-26

Perfect score: 185
Sequence: 1 gsccttaaggccctgcaatc.....tgtatagaataaggagctc 185

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B-COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCITUS-COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	3	US-09-384-305-21
2	185	100.0	185	3	US-09-258-377-26
3	183.4	99.1	185	3	US-09-384-305-22
4	183.4	99.1	185	3	US-09-258-377-27
5	178.6	96.5	185	3	US-09-384-305-20
6	178.6	96.5	185	3	US-09-258-377-25
7	159.6	86.3	10216	2	US-08-875-154-1
8	155.6	84.1	162	1	US-08-434-255-24
9	155.6	84.1	162	1	US-08-459-967-24
10	155.6	84.1	162	1	US-08-460-327-24
11	155.6	84.1	162	1	US-08-459-871-24
12	155.6	84.1	2084	1	US-08-720-899-3
13	155.6	84.1	2084	1	US-08-459-610-3
14	155.6	84.1	2084	2	US-08-343-804-3
15	155.6	84.1	2084	2	US-08-687-399-3
16	155.6	84.1	2084	2	US-08-600-908A-3
17	155.6	84.1	2084	3	US-08-683-838A-3
18	155.6	84.1	2084	3	US-09-182-859-3
19	155.6	84.1	2084	3	US-09-170-670-13
20	155.6	84.1	2084	3	US-09-193-068-31
21	155.6	84.1	2084	3	US-09-183-412-58
22	155.6	84.1	2084	4	US-09-290-734-13
23	155.6	84.1	2084	4	US-09-672-459-3
24	155.6	84.1	2084	4	US-09-636-252A-3
25	155.6	84.1	2084	4	US-09-545-586-13
26	155.6	84.1	2084	4	US-10-186-042-3
27	155.6	84.1	2084	4	US-09-769-864-58

28 155.6 84.1 2604 3 US-09-264-097-3 Sequence 3, Appli
29 155.6 84.1 2604 4 US-09-537-168-5 Sequence 5, Appli
30 44 23.8 44 3 US-09-384-305-15 Sequence 15, Appli
31 44 23.8 44 3 US-09-258-377-19 Sequence 17, Appli
32 39 21.1 42 2 US-08-873-479-37 Sequence 37, Appli
33 39 21.1 42 2 US-09-031-442A-9 Sequence 9, Appli
34 39 21.1 42 2 US-08-972-661A-25 Sequence 25, Appli
35 39 21.1 42 3 US-09-258-377-9 Sequence 9, Appli
36 38.4 20.8 43 2 US-08-873-479-38 Sequence 28, Appli
37 38.4 20.8 43 2 US-09-031-442A-10 Sequence 10, Appli
38 38.4 20.8 43 3 US-08-972-661A-26 Sequence 26, Appli
39 38.4 20.8 43 3 US-09-258-377-10 Sequence 10, Appli
40 32.4 17.5 34 1 US-08-434-255-26 Sequence 26, Appli
41 32.4 17.5 34 1 US-08-459-967-26 Sequence 26, Appli
42 32.4 17.5 34 1 US-08-460-327-26 Sequence 26, Appli
43 32.4 17.5 34 1 US-08-459-871-26 Sequence 26, Appli
44 32.4 17.5 640681 4 US-09-790-988-1 Sequence 1, Appli
45 31.6 17.1 1511 1 US-07-991-867B-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-384-305-21
; Sequence 21, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-21

Query Match 100.0%; Score 185; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.7e-43;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAACCAATAAAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAACCAATAAAATACCTTGT 60
QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAAGGA 120
Db 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAAGGA 120
QY 121 ATAAAGGGGGTGTGACATTTTACTGATATGATATATAATATGTAATAAATG 180
Db 121 ATAAAGGGGGTGTGACATTTTACTGATATGATATATAATATGTAATAAATG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 2
US-09-258-377-26
; Sequence 26, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a

```
/ TITLE OF INVENTION: Bacillus Cell
/ FILE REFERENCE: 5455.200-US
/ CURRENT APPLICATION NUMBER: US/09/258,377
/ CURRENT FILING DATE: 1999-02-26
/ EARLIER APPLICATION NUMBER: 09/031,442
/ EARLIER FILING DATE: 1998-02-26
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 26
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus
US-09-258-377-26

Query Match
Best Local Similarity 100.0%; Score 185; DB 3; Length 185;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
QY 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
US-09-384-305-22
/ Sequence 22, Application US/09384305
/ Patent No. 6184028
/ GENERAL INFORMATION:
/ APPLICANT: Michael D. Thomas
/ APPLICANT: Kimberly M. Brown
/ TITLE OF INVENTION: Polypeptides Having Pectin
/ TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
/ FILE REFERENCE: 5952.000-US
/ CURRENT APPLICATION NUMBER: US/09/384,305
/ CURRENT FILING DATE: 1999-08-26
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 22
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus subtilis
US-09-384-305-22

Query Match
Best Local Similarity 99.1%; Score 183.4; DB 3; Length 185;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
QY 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
US-09-258-377-27
/ Sequence 27, Application US/09258377
/ Patent No. 6255076
/ GENERAL INFORMATION:
/ APPLICANT: Widner, William
/ APPLICANT: Thomas, Michael D.
/ APPLICANT: Sloma, Alan
/ TITLE OF INVENTION: Methods For Producing A polypeptide In a
/ TITLE OF INVENTION: Bacillus Cell
/ FILE REFERENCE: 5455.200-US
/ CURRENT APPLICATION NUMBER: US/09/258,377
/ CURRENT FILING DATE: 1999-02-26
/ EARLIER APPLICATION NUMBER: 09/031,442
/ EARLIER FILING DATE: 1998-02-26
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 27
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus
US-09-258-377-27

Query Match
Best Local Similarity 99.1%; Score 183.4; DB 3; Length 185;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
QY 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
US-09-384-305-20
/ Sequence 20, Application US/09384305
/ Patent No. 6184028
/ GENERAL INFORMATION:
/ APPLICANT: Michael D. Thomas
/ APPLICANT: Kimberly M. Brown
/ TITLE OF INVENTION: Polypeptides Having Pectin
/ TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
/ FILE REFERENCE: 5952.000-US
/ CURRENT APPLICATION NUMBER: US/09/384,305
/ CURRENT FILING DATE: 1999-08-26
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 20
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus subtilis
US-09-384-305-20

Query Match
Best Local Similarity 96.5%; Score 178.6; DB 3; Length 185;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
```

```
/ TITLE OF INVENTION: Bacillus Cell
/ FILE REFERENCE: 5455.200-US
/ CURRENT APPLICATION NUMBER: US/09/258,377
/ CURRENT FILING DATE: 1999-02-26
/ EARLIER APPLICATION NUMBER: 09/031,442
/ EARLIER FILING DATE: 1998-02-26
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 26
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus
US-09-258-377-26

Query Match
Best Local Similarity 100.0%; Score 185; DB 3; Length 185;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
QY 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
US-09-384-305-22
/ Sequence 22, Application US/09384305
/ Patent No. 6184028
/ GENERAL INFORMATION:
/ APPLICANT: Michael D. Thomas
/ APPLICANT: Kimberly M. Brown
/ TITLE OF INVENTION: Polypeptides Having Pectin
/ TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
/ FILE REFERENCE: 5952.000-US
/ CURRENT APPLICATION NUMBER: US/09/384,305
/ CURRENT FILING DATE: 1999-08-26
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 22
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus subtilis
US-09-384-305-22

Query Match
Best Local Similarity 99.1%; Score 183.4; DB 3; Length 185;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
QY 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
US-09-258-377-27
/ Sequence 27, Application US/09258377
/ Patent No. 6255076
/ GENERAL INFORMATION:
/ APPLICANT: Widner, William
/ APPLICANT: Thomas, Michael D.
/ APPLICANT: Sloma, Alan
/ TITLE OF INVENTION: Methods For Producing A polypeptide In a
/ TITLE OF INVENTION: Bacillus Cell
/ FILE REFERENCE: 5455.200-US
/ CURRENT APPLICATION NUMBER: US/09/258,377
/ CURRENT FILING DATE: 1999-02-26
/ EARLIER APPLICATION NUMBER: 09/031,442
/ EARLIER FILING DATE: 1998-02-26
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 27
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus
US-09-258-377-27

Query Match
Best Local Similarity 99.1%; Score 183.4; DB 3; Length 185;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAGAAAAGGA 120
QY 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
US-09-384-305-20
/ Sequence 20, Application US/09384305
/ Patent No. 6184028
/ GENERAL INFORMATION:
/ APPLICANT: Michael D. Thomas
/ APPLICANT: Kimberly M. Brown
/ TITLE OF INVENTION: Polypeptides Having Pectin
/ TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
/ FILE REFERENCE: 5952.000-US
/ CURRENT APPLICATION NUMBER: US/09/384,305
/ CURRENT FILING DATE: 1999-08-26
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 20
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus subtilis
US-09-384-305-20

Query Match
Best Local Similarity 97.8%; Score 178.6; DB 3; Length 185;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
```

Db 1 GGCCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAACACCATATAAAATACCTTGT 60
QY 61 CTGTCAATCAGACAGGTAATTTTATGCTGTCAGACTGTCGCTGTGTGTAATAAAGGA 120
Db 61 CTGTCAATCAGACAGGTAATTTTATGCTGTCAGACTGTCGCTGTGTGTAATAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTTATTTTACTGATATGTAATATATATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTTGATTTATTTTCTGATATGTAATATATATTTGTAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 6

US-09-258-377-25
; Sequence 25, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455,200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; SOFTWARE: FastSeq for Windows Version 3.0
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 25
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-25

Query Match 96.5%; Score 178.6; DB 3; Length 185;

Best Local Similarity 97.8%; Pred. No. 5.8e-41;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAACACCATATAAAATACCTTGT 60
Db 1 GGCCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAACACCATATAAAATACCTTGT 60
QY 61 CTGTCAATCAGACAGGTAATTTTATGCTGTCAGACTGTCGCTGTGTGTAATAAAGGA 120
Db 61 CTGTCAATCAGACAGGTAATTTTATGCTGTCAGACTGTCGCTGTGTGTAATAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTTATTTTACTGATATGTAATATATATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTTGATTTATTTTCTGATATGTAATATATATTTGTAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 7

US-08-875-154-1
; Sequence 1, Application US/08875154
; Patent No. 5882888
; GENERAL INFORMATION:
; APPLICANT: Jorgensen, Steven Truels
; TITLE OF INVENTION: DNA Integration By Transporation
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5882888 No. 5882888disk of No. 5882888th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA

; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,154
; FILING DATE: 17-JUL-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4381.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10216 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "pMOL553"
US-08-875-154-1

Query Match 86.3%; Score 159.6; DB 2; Length 10216;
Best Local Similarity 97.6%; Pred. No. 2.5e-35;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAACACCATATAAAATACCTTGTCTGTCTATCAGACA 73
Db 2237 TGCATCGATTGTTTGAGAAAGAGAACACCATATAAAATACCTTGTCTGTCTATCAGACA 2296
QY 74 GGGTATTTTATGCTCTCCAGACTGTCCTGTGTAATAAAGGAATTAAGGGGGGTT 133
Db 2297 GGGTATTTTATGCTCTCCAGACTGTCCTGTGTAATAAAGGAATTAAGGGGGGTT 2356
QY 134 GACATTTTACTGATATGTAATAAATTTGTAAGAAATG 179
Db 2357 GTTATTATTTTACTGATATGTAATAAATTTGTAAGAAATG 2402

RESULT 8

US-08-434-255-24
; Sequence 24, Application US/08434255
; Patent No. 5621089
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dammann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5621089c No. 5621089disk of No. 5621089th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr., Cheryl H.

REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-434-255-24

Query Match 84.1%; Score 155.6; DB 1; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.4e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGGGT 77
DB 1 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGGGT 60

QY 78 ATTTTATGCTGTCAGACTGTCCTGCTGTGTAATAAATGAGAAATG 179
DB 61 ATTTTATGCTGTCAGACTGTCCTGCTGTGTAATAAATGAGAAATG 162

RESULT 9
US-08-459-967-24
Sequence 24, Application US/08459967
Patent No. 5622841
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Outtrup, Helle
APPLICANT: Dammann, Claus
APPLICANT: Aaslyng, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5622841o No. 5622841disk of No. 5622841th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,967
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-459-967-24

Query Match 84.1%; Score 155.6; DB 1; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.4e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGGGT 77
DB 1 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGGGT 60

QY 78 ATTTTATGCTGTCAGACTGTCCTGCTGTGTAATAAATGAGAAATG 179
DB 61 ATTTTATGCTGTCAGACTGTCCTGCTGTGTAATAAATGAGAAATG 162

RESULT 10
US-08-460-327-24
Sequence 24, Application US/08460327
Patent No. 5622850
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Outtrup, Helle
APPLICANT: Dammann, Claus
APPLICANT: Aaslyng, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5622850o No. 5622850disk of No. 5622850th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,327
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-327-24

Query Match 84.1%; Score 155.6; DB 1; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.4e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGGGT 77
DB 1 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGGGT 60

QY 78 ATTTTATGCTGTCAGACTGTCCTGCTGTGTAATAAATGAGAAATG 179
DB 61 ATTTTATGCTGTCAGACTGTCCTGCTGTGTAATAAATGAGAAATG 162

Db 61 ATTTTATGCTGTCAGACTGCCGCTGTGTAATAAAGGAGGAGGTTGTTA 120
QY 138 TTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 179
Db 121 TTATTTTACTGATATGTAATAATAATTTGTATAAGAAAATG 162

RESULT 11

US-08-459-871-24
; Sequence 24, Application US/08459871
; Patent No. 5650326
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5650326 No. 5650326disk of No. 5650326th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,871
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764.400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-459-871-24

Query Match 84.1%; Score 155.6; DB 1; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.4e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 18 ATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGTCTCATCAGACAGGT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGTCTCATCAGACAGGT 60
QY 78 ATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGAGGAGGTTGACA 137
Db 61 ATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGAGGAGGTTGTTA 120
QY 138 TTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 179
Db 121 TTATTTTACTGATATGTAATAATAATTTGTATAAGAAAATG 162

RESULT 12

US-08-720-899-3
; Sequence 3, Application US/08720899

; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5753460 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowrey Dr. Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 250..1794
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 250..342
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 343..1791
US-08-720-899-3

Query Match 84.1%; Score 155.6; DB 1; Length 2084;
Best Local Similarity 97.5%; Pred. No. 2.3e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 18 ATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGTCTCATCAGACAGGT 77
Db 73 ATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGTCTCATCAGACAGGT 132
QY 78 ATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGAGGAGGTTGACA 137
Db 133 ATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGAGGAGGTTGTTA 192
QY 138 TTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 179
Db 193 TTATTTTACTGATATGTAATAATAATTTGTATAAGAAAATG 234

RESULT 13

US-08-459-610-3
; Sequence 3, Application US/08459610
; Patent No. 5801043

GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thellersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESS: No. 58010430 No. 58010430th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,610
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/343,804
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1794
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-08-459-610-3

Query Match 84.1%; Score 155.6; DB 1; Length 2084;
Best Local Similarity 97.5%; Pred. No. 2.3e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 18 ATCGATTGTTGAGAAAGAGAACCAATAAAATACCTTGTCTGTCTATCAGACAGGTT 77
DB 73 ATCGATTGTTGAGAAAGAGAACCAATAAAATACCTTGTCTGTCTATCAGACAGGTT 132
QY 78 ATTTTATGCTGTCAGACTGCCCTGTGTAAAAAAGGAATAAAGGGGGTTGACA 137
DB 133 ATTTTATGCTGTCAGACTGCCCTGTGTAAAAAAGGAATAAAGGGGGTTGTTA 192
QY 138 TTATTTTACTGATGTATATAATATATTTGTATAAGAAAATG 179
DB 193 TTATTTTACTGATGTATATAATATATTTGTATAAGAAAATG 234

RESULT 14
US-08-343-804-3
Sequence 3, Application US/08343804
Patent No. 5830837
GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thellersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESS: No. 5830837c No. 5830837th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,804
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1794
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-08-343-804-3

GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thellersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESS: No. 5830837c No. 5830837th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,804
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1794
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-08-343-804-3

Query Match 84.1%; Score 155.6; DB 2; Length 2084;
Best Local Similarity 97.5%; Pred. No. 2.3e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 18 ATCGATTGTTGAGAAAGAGAACCAATAAAATACCTTGTCTGTCTATCAGACAGGTT 77
DB 73 ATCGATTGTTGAGAAAGAGAACCAATAAAATACCTTGTCTGTCTATCAGACAGGTT 132
QY 78 ATTTTATGCTGTCAGACTGCCCTGTGTAAAAAAGGAATAAAGGGGGTTGACA 137
DB 133 ATTTTATGCTGTCAGACTGCCCTGTGTAAAAAAGGAATAAAGGGGGTTGTTA 192
QY 138 TTATTTTACTGATGTATATAATATATTTGTATAAGAAAATG 179
DB 193 TTATTTTACTGATGTATATAATATATTTGTATAAGAAAATG 234

RESULT 15
US-08-687-399-3
Sequence 3, Application US/08687399
Patent No. 5928381
GENERAL INFORMATION:
APPLICANT: Toit, Annette H.
APPLICANT: Marcher, Dorthé
APPLICANT: Pedersen, Hanne H.
APPLICANT: Nilsson, Thomas E.


```

; TITLE OF INVENTION: A Combined Desizing and Bleaching
; TITLE OF INVENTION: Process
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 59283810 No. 59283810disk of No. 59283810th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,399
; FILING DATE:
; CLASSIFICATION: 008
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4127,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 250..342
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 250..1791
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 343..1791
; US-08-687-399-3

```

```

Query Match      84.1%; Score 155.6; DB 2; Length 2084;
Best Local Similarity 97.5%; Pred. No. 2.3e-34;
Matches 159; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAGAGAGACCATATAACCTTGTCTGTCATCAGACAGGCT 77
   |||||
Db 73 ATCGATTGTTTGAGAAAGAGAGACCATATAACCTTGTCTGTCATCAGACAGGCT 132
   |||||

Qy 78 ATTTTATGCTGTCAGACTGCCGTGTGTAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 137
   |||||
Db 133 ATTTTATGCTGTCAGACTGCCGTGTGTAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 192
   |||||

Qy 138 TTATTTTACTGATGATATATATATTTGTTAAGAAATG 179
   |||||
Db 193 TTATTTTACTGATGATATATATTTGTTAAGAAATG 234
   |||||

```

Search completed: March 19, 2004, 03:17:58
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 00:21:10 ; Search time 215 Seconds
(without alignments)
3655.426 Million cell updates/sec

Title: US-09-834-271A-26
Perfect score: 185
Sequence: 1 ggccttaaggcctgaatc.....tgtataagaaagagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124093041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	100.0	185	2	Aaz23323 B. amylol
2	185	100.0	185	4	Aaf62646 Consensus
3	185	100.0	185	4	Aad09923 Bacillus
4	183.4	99.1	185	2	Aaz23324 B. amylol
5	183.4	99.1	185	4	Aaf62647 Consensus
6	183.4	99.1	185	4	Aad09924 Bacillus
7	180.2	97.4	185	2	Aaz23325 Bacillus
8	180.2	97.4	185	4	Aad09931 Bacillus
9	178.6	96.5	185	2	Aaz23322 B. thurin
10	178.6	96.5	185	4	Aaf62645 Wild type
11	178.6	96.5	185	4	Aad09922 Bacillus
12	159.6	86.3	2166	6	Aad34433 C-termina
13	159.6	86.3	2166	6	Abi40498 DNA const
14	159.6	86.3	2267	6	Aad34434 Subtilisi
15	159.6	86.3	2267	6	Abi40499 DNA const
16	159.6	86.3	2588	6	Aad34436 PamyL ATG
17	159.6	86.3	2588	6	Abi40501 DNA const
18	156.4	84.5	10216	2	Aat39279 Transposo
19	156.4	84.1	162	2	Aat85631 BSN Promo
20	155.6	84.1	249	3	Aaa3677 Bacillus
21	155.6	84.1	249	7	Acag1514 Modified
22	155.6	84.1	270	3	Aaa3678 Modified
23	155.6	84.1	537	2	Aaa22579 Sequence

24	155.6	84.1	576	1	AAN20046
25	155.6	84.1	2084	2	AAQ88067 Bacillus
26	155.6	84.1	2084	2	AAQ95032 Bacillus
27	155.6	84.1	2084	2	AAX59681 DNA encod
28	155.6	84.1	2084	2	AAX57596 Wild type
29	155.6	84.1	2084	3	AAA48484 Bacillus
30	155.6	84.1	2084	3	ADL50568 B. amylol
31	155.6	84.1	2084	6	AAS20026 Bacillus
32	155.6	84.1	2084	6	AAI72215 Bacillus
33	155.6	84.1	2084	6	ABL96211 Termamyl-
34	155.6	84.1	2604	2	AZ21079 Bacillus
35	155.6	84.1	2604	4	AAA37850
36	155.6	84.1	5661	7	AD49392 PMOL995 P
37	143.6	77.6	2083	2	AAV02472 Bacillus
38	95.2	51.5	13222	4	AAD25109 Nucleotid
39	88.6	47.9	6637	6	AAH29902 Plasmid p
40	48	25.9	48	4	AAD09917 Oligo #2
41	44	23.8	44	2	AZ23340 Bacillus
42	44	23.8	44	4	AAF62640 amyo prom
43	44	23.8	44	4	AAD09916 Oligo #1
44	39	21.1	42	2	AAV38601 PCR prime
45	39	21.1	42	2	AZ23330 Bacillus

ALIGNMENTS

RESULT 1
AAZ23323
ID AAZ23323 standard; DNA; 185 BP.
XX
AC AAZ23323;
DT 06-DEC-1999. (first entry)
XX
DE B. amylolifaciens amyo promoter DNA #1.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone; amyo;
KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
XX
OS Bacillus amylolifaciens.
XX
PN WO9943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US004360.
XX
PR 26-FEB-1999; 99US-00031442.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
DR WPI; 1999-561370/47.
XX
PT Production of polypeptide in Bacillus using specific promoters,
XX particularly for producing enzymes.
XX
PS Claim 44; Page 90; 90pp; English.
XX
CC This invention describes a novel method for the production of a
CC polypeptide in Bacillus using specific tandem or consensus promoters. The
CC method is used to produce homologous or particularly heterologous
CC proteins, particularly enzymes (specifically serine proteases, maltogenic
CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
CC etc. The specified promoters provide increased expression of the sequence
CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents a Bacillus amylolifaciens
CC alpha-amylase amyo promoter which is used to describe the method of the
CC invention

applicant

CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a *Bacillus* strain. The present
 CC sequence is *Bacillus amyloliquefaciens* alpha-amylase (amyQ) consensus
 CC mutated promoter, used in the exemplification of the invention
 XX
 SQ Sequence 185 BP; 63 A; 25 C; 42 G; 55 T; 0 U; 0 Other;

Query Match 100.0%; Score 185; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 1.6e-41;
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
 DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
 QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAAGGA 120
 DB 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAAGGA 120
 QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAAGAAAATGG 180
 DB 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAAGAAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 4
 AAZ23324
 ID AAZ23324 standard; DNA; 185 BP.

XX AAZ23324;

XX 06-DEC-1999 (first entry)

XX *B. amyloliquefaciens* amyQ promoter DNA #2.

XX Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
 XX antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.

XX *Bacillus amyloliquefaciens*.

XX WO9943835-A2.

XX 02-SEP-1999.

XX 26-FEB-1999; 99WO-US004360.

XX 26-FEB-1998; 98US-00031442.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Widner W, Sloma A, Thomas MD;

XX WPI; 1999-561370/47.

XX Production of polypeptide in *Bacillus* using specific promoters,
 XX particularly for producing enzymes.

XX Claim 44; Page 90; 90pp; English.

XX This invention describes a novel method for the production of a
 CC polypeptide in *Bacillus* using specific tandem or consensus promoters. The
 CC method is used to produce homologous or particularly heterologous
 CC proteins, particularly enzymes (specifically serine protease, maltogenic
 CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
 CC etc. The specified promoters provide increased expression of the sequence
 CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be
 CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents a *Bacillus amyloliquefaciens*

CC alpha-amylase amyQ promoter which is used to describe the method of the
 CC invention
 XX
 SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;

Query Match 99.1%; Score 183.4; DB 2; Length 185;
 Best Local Similarity 99.5%; Pred. No. 4.4e-41;
 Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
 DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
 QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAAGGA 120
 DB 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAAGGA 120
 QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAAGAAAATGG 180
 DB 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAAGAAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 5
 AAF62647
 ID AAF62647 standard; DNA; 185 BP.

XX AAF62647;

XX 03-MAY-2001 (first entry)

XX Consensus amyQ promoter #2.

XX Pectin acetyltransferase; degrade; plant cell wall; ss.

XX Unidentified.

XX US6184028-B1.

XX 06-FEB-2001.

XX 26-AUG-1999; 99US-00384305.

XX 26-AUG-1999; 99US-00384305.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Thomas MD, Brown KW;

XX WPI; 2001-190946/19.

XX Novel isolated polypeptide having pectin acetyltransferase activity useful
 XX for degrading pectic substances and in degradation or modification of
 XX acetylated pectins and plant cell walls.

XX Example 9; Col 43-44; 35pp; English.

XX The present invention relates to *Bacillus subtilis* pectin acetyltransferase
 CC protein. The invention is useful for degrading a pectic substance. It is
 CC also useful for degrading soluble and insoluble pectins with varying
 CC degrees of esterification, clarification etc. The protein may be used
 CC alone or in combination with other enzymes for the degradation or
 CC modification of acetylated pectins, degradation or modification of plant
 CC cell walls

XX Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;

Query Match 99.1%; Score 183.4; DB 4; Length 185;
 Best Local Similarity 99.5%; Pred. No. 4.4e-41;
 Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACACATATAAAATACCTTGT 60
 DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACACATATAAAATACCTTGT 60
 QY 61 CTGTCAATCAGACAGAGGTATTTTATGCTGTCACAGCTGTCGGCTGTGTAATAAAAGGA 120
 DB 61 CTGTCAATCAGACAGAGGTATTTTATGCTGTCACAGCTGTCGGCTGTGTAATAAAAGGA 120
 QY 121 ATAAAGGGGGTGCATATTTTACTGATATGATATATAATTTGTATAGAAATGG 180
 DB 121 ATAAAGGGGGTGCATATTTTACTGATATGATATATAATTTGTATAGAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 6

AAD09924
 ID AAD09924 standard; DNA; 185 BP.

AC AAD09924;

DT 12-SEP-2001 (first entry)

DE Bacillus amyloliquefaciens alpha-amylase (amyQ) consensus promoter #2.

XX Bacillus cell; tandem promoter; consensus promoter; mutant;
 KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
 KW cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.
 XX
 OS Bacillus amyloliquefaciens.
 OS Synthetic.

XX Key Location/Qualifiers
 FH replace(135, T)
 FT /*tag= a
 FT mutation
 FT replace(136, T)
 FT /*tag= b
 FT mutation
 FT replace(156, A)
 FT /*tag= c

FT /note= "In column 26, a T to A mutation was said to occur
 at position 116 but no such mutation was found in the
 sequence shown in the sequence listing of the
 specification"

PN US6255076-B1.

XX 03-JUL-2001.

XX 26-FEB-1999; 99US-00258377.

XX 26-FEB-1998; 98US-00031442.

XX (NOVO) NOVOZYMES BIOTECH INC.

XX Widner W, Sloma A, Thomas MD;

XX WPI; 2001-440518/47.

XX Producing a polypeptide in a Bacillus strain comprises cultivating a
 PT Bacillus cell, which contains a nucleic acid construct comprising a
 PT tandem promoter and consensus promoters.

XX Example 18; Col 43-44; 54pp; English.

XX The present invention relates to a method for producing a polypeptide,
 CC comprises cultivating a Bacillus cell, which contains a nucleic acid
 CC construct comprising a tandem promoter and consensus promoters. The
 CC Bacillus is cultivated in a medium conducive for the production of the
 CC polypeptide. Each promoter sequence of the tandem promoter is operably
 CC linked to a nucleic acid sequence encoding the polypeptide. The tandem

CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
 CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
 CC subtilisin Carlsberg gene promoter have the sequence TTGACA for the -35
 CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a Bacillus strain. The present
 CC sequence is Bacillus amyloliquefaciens alpha-amylase (amyQ) consensus
 CC mutated promoter, used in the exemplification of the invention
 XX

SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;

Query Match 99.1%; Score 183.4; DB 4; Length 185;

Best Local Similarity 99.5%; Pred. No. 4.4e-41;

Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACACATATAAAATACCTTGT 60
 DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACACATATAAAATACCTTGT 60

QY 61 CTGTCAATCAGACAGAGGTATTTTATGCTGTCACAGCTGTCGGCTGTGTAATAAAAGGA 120

DB 61 CTGTCAATCAGACAGAGGTATTTTATGCTGTCACAGCTGTCGGCTGTGTAATAAAAGGA 120

QY 121 ATAAAGGGGGTGCATATTTTACTGATATGATATATAATTTGTATAGAAATGG 180

DB 121 ATAAAGGGGGTGCATATTTTACTGATATGATATATAATTTGTATAGAAATGG 180

QY 181 AGCTC 185

DB 181 AGCTC 185

RESULT 7

AAZ23325
 ID AAZ23325 standard; DNA; 185 BP.

XX AC AAZ23325;

DT 06-DEC-1999 (first entry)

DE Bacillus sp. amyQ promoter consensus DNA.

XX Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;

KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.

XX Bacillus sp.

XX WO9943835-A2.

XX 02-SEP-1999.

XX 26-FEB-1999; 99WO-US004360.

XX 26-FEB-1998; 98US-00031442.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Widner W, Sloma A, Thomas MD;

XX WPI; 1999-561370/47.

XX Production of polypeptide in Bacillus using specific promoters,
 PT particularly for producing enzymes.

XX Disclosure; Fig 21; 90pp; English.

XX This invention describes a novel method for the production of a
 CC polypeptide in Bacillus using specific tandem or consensus promoters. The
 CC method is used to produce homologous or particularly heterologous
 CC proteins, particularly enzymes (specifically serine protease, maltogenic
 CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
 CC etc. The specified promoters provide increased expression of the sequence

CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be
 CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents the consensus sequence of a
 CC Bacillus sp. alpha-amylase amyQ promoter which is used to describe the
 CC method of the invention
 XX
 SQ Sequence 185 BP; 63 A; 24 C; 42 G; 56 T; 0 U; 0 Other;
 Query Match 97.4%; Score 180.2; DB 2; Length 185;
 Best Local Similarity 98.4%; Pred. No. 3.3e-40;
 Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
 DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
 QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
 DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
 QY 121 ATAAAGGGGGTTCACATTTATTTTACTGATATGTAATAATATTTGTAAGAAATGG 180
 DB 121 ATAAAGGGGGTTCACATTTATTTTACTGATATGTAATAATATTTGTAAGAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185
 RESULT 8
 AAD09931
 ID AAD09931 standard; DNA; 185 BP.
 AC AAD09931;
 DT 12-SEP-2001 (first entry)
 XX Bacillus amyloliquefaciens amyQ promoter, alternative version.
 DE
 XX Bacillus cell; tandem promoter; consensus promoter;
 KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
 KW cryIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.
 XX
 OS Bacillus amyloliquefaciens.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 116
 FT /*tag= a
 FT /note= "Base T is found at this location in the sequence
 FT shown in the sequence listing of the specification
 FT (AAD09922)."
 XX
 PN US6255076-B1.
 XX
 PD 03-JUL-2001.
 XX
 PF 26-FEB-1999; 99US-00258377.
 XX
 PR 26-FEB-1998; 98US-00031442.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 XX
 PI Widner W, Sloma A, Thomas MD;
 XX
 DR WPI; 2001-440518/47.
 XX
 XX Producing a polypeptide in a Bacillus strain comprises cultivating a
 PT Bacillus cell, which contains a nucleic acid construct comprising a
 PT tandem promoter and consensus promoters.
 XX
 PS Example 18; Fig 21; 54pp; English.
 XX

CC The present invention relates to a method for producing a polypeptide,
 CC comprises cultivating a Bacillus cell, which contains a nucleic acid
 CC construct comprising a tandem promoter and consensus promoters. The
 CC Bacillus is cultivated in a medium conducive for the production of the
 CC polypeptide. Each promoter sequence of the tandem promoter is operably
 CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
 CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIA
 CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIA promoter or
 CC subtilisin Carlsberg gene promoter have the sequence TTGACA for the -35
 CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a Bacillus strain. The present
 CC sequence is an alternative of Bacillus amyloliquefaciens alpha-amylase
 CC (amyQ) promoter, used in the exemplification of the invention. Note: This
 CC sequence is stated as being the same as that shown as SEQ ID NO 25
 CC (AAD09922) in column 43-44 of the specification. However the sequences
 CC differ by at position 116
 XX
 SQ Sequence 185 BP; 63 A; 24 C; 42 G; 56 T; 0 U; 0 Other;
 Query Match 97.4%; Score 180.2; DB 4; Length 185;
 Best Local Similarity 98.4%; Pred. No. 3.3e-40;
 Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
 DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
 QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
 DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
 QY 121 ATAAAGGGGGTTCACATTTATTTTACTGATATGTAATAATATTTGTAAGAAATGG 180
 DB 121 ATAAAGGGGGTTCATTTATTTTACTGATATGTAATAATATTTGTAAGAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185
 RESULT 9
 AAZ23322
 ID AAZ23322 standard; DNA; 185 BP.
 AC AAZ23322;
 DT 06-DEC-1999 (first entry)
 DE B. thuringiensis wild-type amyQ promoter DNA.
 XX
 XX Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
 KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
 XX
 OS Bacillus thuringiensis.
 XX
 PN WO9943835-A2.
 XX
 PD 02-SEP-1999.
 XX
 PF 26-FEB-1999; 99WO-US004360.
 XX
 PR 26-FEB-1998; 98US-00031442.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX
 PI Widner W, Sloma A, Thomas MD;
 XX
 DR WPI; 1999-561370/47.
 XX
 XX Production of polypeptide in Bacillus using specific promoters,
 PT particularly for producing enzymes.

```
XX Disclosure; Page 90; 90pp; English.
PS
XX
CC This invention describes a novel method for the production of a
CC polypeptide in Bacillus using specific tandem or consensus promoters. The
CC method is used to produce homologous or particularly serine protease, maltogenic
CC proteins, particularly enzymes (specifically serine protease, maltogenic
CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
CC etc. The specified promoters provide increased expression of the sequence
CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents a Bacillus thuringiensis
CC alpha-amylase amyQ promoter which is used to describe the method of the
CC invention
XX
SQ Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;
Query Match 96.5%; Score 178.6; DB 2; Length 185;
Best Local Similarity 97.8%; Pred. No. 9.2e-40;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
DB 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
QY 121 ATAAAGGGGGTTCACATTTATTTTACTGATATGTAATAATATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCATTTATTTTACTGATATGTAATAATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185
RESULT 10
ID AAF62645 standard; DNA; 185 BP.
XX
AC AAF62645;
XX
DT 03-MAY-2001 (first entry)
XX
DE Wild type amyQ promoter.
XX
KW Pectin acetyltransferase; degrade; plant cell wall; ss.
XX
OS Unidentified.
XX
PN US6184028-B1.
XX
PD 06-FEB-2001.
XX
PF 26-AUG-1999; 99US-00384305.
XX
PR 26-AUG-1999; 99US-00384305.
XX
PA (NOVO ) NOVO NORDISK BIOTECH INC.
XX
PI Thomas MD, Brown KM;
XX
DR WPI; 2001-190946/19.
XX
Novel isolated polypeptide having pectin acetyltransferase activity useful
PT for degrading pectic substances and in degradation or modification of
PT acetylated pectins and plant cell walls.
XX
PS Example 9; Fig 9; 35pp; English.
XX
```

```
CC The present invention relates to Bacillus subtilis pectin acetyltransferase
CC protein. The invention is useful for degrading a pectic substance. It is
CC also useful for degrading soluble and insoluble pectins with varying
CC degrees of esterification, clarification etc. The protein may be used
CC alone or in combination with other enzymes for the degradation or
CC modification of acetylated pectins, degradation or modification of plant
CC cell walls
XX
SQ Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;
Query Match 96.5%; Score 178.6; DB 4; Length 185;
Best Local Similarity 97.8%; Pred. No. 9.2e-40;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
DB 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
QY 121 ATAAAGGGGGTTCACATTTATTTTACTGATATGTAATAATATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCATTTATTTTACTGATATGTAATAATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185
RESULT 11
ID AAD09922 standard; DNA; 185 BP.
XX
AC AAD09922;
XX
DT 12-SEP-2001 (first entry)
XX
DE Bacillus amyloliquefaciens alpha-amylase (amyQ) promoter.
XX
KW Bacillus cell; tandem promoter; consensus promoter;
KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
KW cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.
XX
OS Bacillus amyloliquefaciens.
XX
PN US6255076-B1.
XX
PD 03-JUL-2001.
XX
PF 26-FEB-1999; 99US-00258377.
XX
PR 26-FEB-1998; 98US-00031442.
XX
PA (NOVO ) NOVOZYMES BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
DR WPI; 2001-440518/47.
XX
Producing a polypeptide in a Bacillus strain comprises cultivating a
PT Bacillus cell, which contains a nucleic acid construct comprising a
PT tandem promoter and consensus promoters.
XX
PS Example 18; Col 43-44; 54pp; English.
XX
The present invention relates to a method for producing a polypeptide,
CC comprises cultivating a Bacillus cell, which contains a nucleic acid
CC construct comprising a tandem promoter and consensus promoters. The
CC Bacillus is cultivated in a medium conducive for the production of the
CC polypeptide. Each promoter sequence of the tandem promoter is operably
CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
```

CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
 CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
 CC subtilisin Carlsberg gene promoter have the sequence TTGACA for the -35
 CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a Bacillus strain. The present
 CC sequence is Bacillus anoliquefaciens alpha-amylase (amyQ) promoter,
 CC used in the exemplification of the invention. Note: This sequence is
 CC stated as being the same as that shown as SEQ ID NO 25 (AAB0931) in
 CC figure 21 of the specification. However the sequences differ by at
 CC position 116
 XX
 SQ Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;

Query Match 96.5%; Score 178.6; DB 4; Length 185;
 Best Local Similarity 97.8%; Pred. No. 9.2e-40;
 Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGCCCTGCAATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGT 60
 DB 1 GGCCTTAAGGCCCTGCAATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGT 60
 QY 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCAGACTGTCGCGTGTGTAAGAAAGGA 120
 DB 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCAGACTGTCGCGTGTGTAAGAAATAGGA 120
 QY 121 ATAAAGGGGGTTCACATTTATTTTACTGATGATGATATATATTTGTAAGAAATAGG 180
 DB 121 ATAAAGGGGGTTCATTTATTTTACTGATGATGATGATATATTTGTAAGAAATAGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 12
 AAD34433
 ID AAD34433 standard; DNA; 2166 BP.
 XX
 AC AAD34433;
 XX
 DT 29-AUG-2003 (revised)
 DT 16-JUL-2002 (first entry)
 DE C-terminal Subtilisin 309-CI-2A fusion protein encoding construct A DNA.
 KW Protease; screening; barley; chymotrypsin inhibitor; CI-2A; laundry;
 KW detergent industry; dishwash; fabric softener; subtilisin 309; savinase;
 KW chimeric; ds.
 XX
 OS Hordeum sp.
 OS Bacillus lentus.
 OS Chimeric.
 XX
 PN WO200218588-A1.
 XX
 PD 07-VAR-2002.
 XX
 XX 17-JUL-2001; 2001WO-DK000503.
 XX
 XX 29-AUG-2000; 2000DK-00001273.
 PR 05-SEP-2000; 2000US-0230017P.
 XX
 PA (NOVO) NOVOZYMES AS.
 XX
 XX Pedersen PE, Norregaard-Madsen M;
 PI
 XX WPI; 2002-329776/36.
 XX
 XX Screening protease/protease inhibitor gene library for gene encoding the
 PT same, by constructing host cell having the gene, cultivating cell,
 PT dissociating complex formed between them, selecting protein and isolating

PT gene.
 PS Example 1; Page 47-48; 54pp; English.
 XX
 CC The invention relates to a method for screening a protease/protease
 CC inhibitor gene library for a gene encoding a protease of interest. The
 CC method comprises introducing a first gene of the protease gene library
 CC and a second gene encoding a protease inhibitor into a host cell;
 CC culturing the host cell, wherein the cell expresses the first and the
 CC second genes to produce a complex of a protease and the inhibitor;
 CC dissociating the inhibitor from the complex and selecting the protease of
 CC interest and isolating the encoding gene. A complex may be used for a
 CC number of industrial applications, in particular within the detergent
 CC industry, thus, relates to a cleaning or detergent composition,
 CC preferably a laundry or dishwash composition comprising the complex,
 CC where hand or machine laundry detergent composition including a laundry
 CC additive composition suitable for pre-treatment of stained fabrics and a
 CC rinse added composition, or be formulated as a detergent composition for
 CC use in general household hard surface cleaning operations or be
 CC formulated for hand or machine dish washing operations. The present
 CC sequence is a construct encoding a fusion protein containing C-terminal
 CC region of subtilisin 309 (savinase) and barley chymotrypsin inhibitor CI-
 CC 2A protein used in the invention. (Updated on 29-AUG-2003 to standardise
 CC OS field)
 XX
 SQ Sequence 2166 BP; 610 A; 442 C; 582 G; 532 T; 0 U; 0 Other;

Query Match 86.3%; Score 159.6; DB 6; Length 2166;
 Best Local Similarity 97.6%; Pred. No. 2.8e-34;
 Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 14 TCCAATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGTCTGTCATCAGACA 73
 DB 172 TCCAATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGTCTGTCATCAGACA 231
 QY 74 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAAGGAATAAAGGGGGTT 133
 DB 232 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAAAGGGGGTT 291
 QY 134 GACATTTTACTGATGATGATATATATTTGTAAGAAATG 179
 DB 292 GTTATTATTACTGATGATGTAATAATTTGTAAGAAATG 337

RESULT 13
 ABL40498
 ID ABL40498 standard; DNA; 2166 BP.
 XX
 AC ABL40498;
 XX
 DT 17-JUN-2002 (first entry)
 XX
 DE DNA construct A comprising Savinase-CI-2A fusion sequence.
 XX
 KW Protease; allergenicity; detergent; subtilase; S11; S12; additive;
 KW protease inhibitor; barley; chymotrypsin; CI-2A; fusion protein;
 KW subtilisin; savinase; apr gene; ds.
 XX
 OS Synthetic.
 OS Bacillus sp.
 OS Hordeum sp.
 XX
 XX WO200216619-A1.
 XX
 XX 28-FEB-2002.
 XX
 XX 09-JUL-2001; 2001WO-DK000479.
 PF
 XX 21-AUG-2000; 2000DK-00001233.
 PR 24-AUG-2000; 2000US-0227501P.
 XX
 XX (NOVO) NOVOZYMES AS.
 PA
 XX


```

PI Pedersen PE;
XX WPI; 2002-242031/29.
XX Producing protease-inhibitor compositions with reduced allergenicity,
XX stability and activity for use in detergent compositions.
XX Example 1; Page 45-46; 52pp; English.
XX The invention provides a method for producing a protease-inhibitor
XX complex with reduced allergenicity for use in detergents. The method
XX involves (a) constructing a fusion polynucleotide sequence in frame, the
XX sequence comprising a first gene encoding a protease (preferably a
XX subtilase S11 or S12) and a second gene encoding a protease inhibitor;
XX (b) introducing the sequence into a host cell; and (c) cultivating the
XX host cell (the host cell expresses the sequence and produces a non-
XX covalently linked complex of the protease and the inhibitor). The
XX protease-inhibitor complex produced by the method may be used as an
XX nucleotide sequence of a DNA construct comprising the savinase-CI-2A
XX fusion sequence
XX Sequence 2166 BP; 610 A; 442 C; 582 G; 532 T; 0 U; 0 Other;
XX
XX Query Match 86.3%; Score 159.6; DB 6; Length 2166;
XX Best Local Similarity 97.6%; Pred. No. 2.8e-34;
XX Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 14 TGCATCGATTGTTGAGAAAGAGAGACCATAAATACCTTGCTGTCATCAGACA 73
XX DB 172 TGCATCGATTGTTGAGAAAGAGAGACCATAAATACCTTGCTGTCATCAGACA 231
XX
XX QY 74 GGGTATTTTATGCTGTCAGACTGCGCTGTGTAATAAAGGAATAAGGGGGTT 133
XX DB 232 GGGTATTTTATGCTGTCAGACTGCGCTGTGTAATAAAGGAATAAGGGGGTT 291
XX
XX QY 134 GACATTTATTTTACTGATGTAATATATATTTGTAAGAAATG 179
XX DB 292 GTTATTTATTTTACTGATGTAATATATATTTGTAAGAAATG 337
XX
XX RESULT 14
XX AAD34434
XX ID AAD34434 standard; DNA; 2267 BP.
XX AC AAD34434;
XX DT 29-AUG-2003 (revised)
XX DT 16-JUL-2002 (first entry)
XX DE Subtilisin 309-CI-2A transcriptional product encoding construct B DNA.
XX KW Protease; screening; barley; chymotrypsin inhibitor; CI-2A; laundry;
XX KW detergent industry; dishwash; fabric softener; subtilisin 309; savinase;
XX KW chimeric; ds.
XX OS Hordeum sp.
XX OS Bacillus lentus.
XX OS Chimeric.
XX PN WO200218588-A1.
XX XX 07-MAR-2002.
XX XX 17-JUL-2001; 2001WO-DK000503.
XX PF 29-AUG-2000; 2000DK-00001273.
XX PR 05-SEP-2000; 2000US-0230017P.
XX XX (NOVO ) NOVOZYMES AS.
XX PA Pedersen PE, Norregaard-Madsen M;
XX PI WPI; 2002-329776/36.
XX DR

```

```

XX Screening protease/protease inhibitor gene library for gene encoding the
XX same, by constructing host cell having the gene, cultivating cell,
XX dissociating complex formed between them, selecting protein and isolating
XX gene.
XX Example 1; Page 48; 54pp; English.
XX The invention relates to a method for screening a protease/protease
XX inhibitor gene library for a gene encoding a protease of interest. The
XX method comprises introducing a first gene of the protease gene library
XX and a second gene encoding a protease inhibitor into a host cell;
XX cultivating the host cell, wherein the cell expresses the first and the
XX second genes to produce a complex of a protease and the inhibitor;
XX dissociating the inhibitor from the complex and selecting the protease of
XX interest and isolating the encoding gene. A complex may be used for a
XX number of industrial applications, in particular within the detergent
XX industry, thus, relates to a cleaning or detergent composition,
XX preferably a laundry or dishwash composition comprising the complex,
XX where hand or machine laundry detergent composition including a laundry
XX additive composition suitable for pre-treatment of stained fabrics and a
XX rinse added composition, or be formulated as a detergent composition for
XX use in general household hard surface cleaning operations or be
XX formulated for hand or machine dish washing operations. The present
XX sequence is a construct encoding a transcriptional fusion product
XX containing C-terminal region of subtilisin 309 (savinase) and barley
XX chymotrypsin inhibitor CI-2A protein used in the invention. (Updated on
XX 29-AUG-2003 to standardise OS field)
XX SQ Sequence 2267 BP; 640 A; 462 C; 601 G; 564 T; 0 U; 0 Other;
XX
XX Query Match 86.3%; Score 159.6; DB 6; Length 2267;
XX Best Local Similarity 97.6%; Pred. No. 2.8e-34;
XX Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 14 TGCATCGATTGTTGAGAAAGAGAGACCATAAATACCTTGCTGTCATCAGACA 73
XX DB 172 TGCATCGATTGTTGAGAAAGAGAGACCATAAATACCTTGCTGTCATCAGACA 231
XX
XX QY 74 GGGTATTTTATGCTGTCAGACTGCGCTGTGTAATAAAGGAATAAGGGGGTT 133
XX DB 232 GGGTATTTTATGCTGTCAGACTGCGCTGTGTAATAAAGGAATAAGGGGGTT 291
XX
XX QY 134 GACATTTATTTTACTGATGTAATATATATTTGTAAGAAATG 179
XX DB 292 GTTATTTATTTTACTGATGTAATATATATTTGTAAGAAATG 337
XX
XX RESULT 15
XX ABL40499
XX ID ABL40499 standard; DNA; 2267 BP.
XX XX
XX AC ABL40499;
XX XX
XX DT 17-JUN-2002 (first entry)
XX DE DNA construct B comprising Savinase-CI-2A fusion sequence.
XX DE
XX KW Protease; allergenicity; detergent; subtilase; S11; S12; additive;
XX KW protease inhibitor; barley; chymotrypsin; CI-2A; fusion protein;
XX KW subtilisin; savinase; apr gene; ds.
XX OS Synthetic.
XX OS Bacillus sp.
XX OS Hordeum sp.
XX XX
XX PN WO200216619-A1.
XX XX
XX PD 28-FEB-2002.
XX XX
XX PF 09-JUL-2001; 2001WO-DK000479.
XX XX
XX PR 21-AUG-2000; 2000DK-00001233.
XX PR

```

PR 24-AUG-2000; 2000US-0227501P.
XX PA (NOVO) NOVOZYMES AS.
XX PI Pedersen PE;
XX DR WPI; 2002-242031/29.
XX PT Producing protease-inhibitor compositions with reduced allergenicity,
XX PT stability and activity for use in detergent compositions.
XX PS Example 1; Page 46; 52pp; English.
XX CC The invention provides a method for producing a protease-inhibitor
XX CC complex with reduced allergenicity for use in detergents. The method
XX CC involves (a) constructing a fusion polynucleotide sequence in frame, the
XX CC sequence comprising a first gene encoding a protease (preferably a
XX CC subtilase S11 or S12) and a second gene encoding a protease inhibitor;
XX CC (b) introducing the sequence into a host cell; and (c) cultivating the
XX CC host cell (the host cell expresses the sequence and produces a non-
XX CC covalently linked complex of the protease and the inhibitor). The
XX CC protease-inhibitor complex produced by the method may be used as an
XX CC nucleotide sequence of a DNA construct comprising the savinase-CI-2A
XX CC fusion sequence
XX SQ Sequence 2267 BP; 640 A; 462 C; 601 G; 564 T; 0 U; 0 Other;
Query Match 86.3%; Score 159.6; DB 6; Length 2267;
Best Local Similarity 97.6%; Pred. No. 2.8e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 14 TGCATCGATGTTTGAGAAAGAGAGAGACCATTAATAATACCTTCTCTCATCAGACA 73
Db 172 TGCATCGATGTTTGAGAAAGAGAGAGACCATTAATAATACCTTCTCTCATCAGACA 231
QY 74 GGGTATTTTATGCTCTCCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTT 133
Db 232 GGGTATTTTATGCTCTCCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTT 291
QY 134 GACATATTTTACTGATATGATATATAATATTTTGTATAGAAAATG 179
Db 292 GTTATTTTACTGATATGATATATAATATTTTGTATAGAAAATG 337

Search completed: March 19, 2004, 01:38:17
Job time : 218 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 00:23:20 ; Search time 1078.5 Seconds
(without alignments)
7434.827 Million cell updates/sec

Title: US-09-834-271A-26
Perfect score: 185
Sequence: 1 ggccttaaggccgtgcattc.....tgataaagaatgagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.em.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pin.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	100.0	185	6	AR160502 Sequence
2	185	100.0	185	6	AX088920 Sequence
3	185	100.0	185	6	BD136798 Process f
4	183.4	99.1	185	6	AR160503 Sequence
5	183.4	99.1	185	6	AX088921 Sequence
6	183.4	99.1	185	6	BD136799 Process f
7	178.6	96.5	185	6	AR160501 Sequence
8	178.6	96.5	185	6	AX088919 Sequence
9	178.6	96.5	185	6	BD136797 Process f
10	155.6	84.1	162	6	I40596 Sequence 24
11	155.6	84.1	162	6	I40867 Sequence 24
12	155.6	84.1	162	6	I40895 Sequence 24
13	155.6	84.1	162	6	I56837 Sequence 24
14	155.6	84.1	249	6	BD174625 Modified
15	155.6	84.1	537	1	BAAMYL
16	155.6	84.1	576	6	A00155 Nucleotide
17	155.6	84.1	576	6	A00156 Nucleotide
18	155.6	84.1	576	6	A00157 Nucleotide
19	155.6	84.1	576	6	A00158 Nucleotide
20	155.6	84.1	576	6	A00604 B.amyloliqu
21	155.6	84.1	576	6	A00605 B.amyloliqu
22	155.6	84.1	2084	1	BACAAAM
23	155.6	84.1	2084	6	AR008285 Sequence
24	155.6	84.1	2084	6	AR037275 Sequence
25	155.6	84.1	2084	6	AR052145 Sequence
26	155.6	84.1	2084	6	AR087551 Sequence
27	155.6	84.1	2084	6	AR129916 Sequence
28	155.6	84.1	2084	6	AR137905 Sequence
29	155.6	84.1	2084	6	AR143262 Sequence
30	155.6	84.1	2084	6	BD230131 Alpha-amy
31	155.6	84.1	2084	6	AR224261 Sequence
32	155.6	84.1	2084	6	AR285306 Sequence
33	155.6	84.1	2084	6	AR428861 Sequence
34	155.6	84.1	2084	6	AX244195 Sequence
35	155.6	84.1	2084	6	AX305020 Sequence
36	155.6	84.1	2084	6	AX339268 Sequence
37	155.6	84.1	2084	6	AX370717 Sequence
38	155.6	84.1	2084	6	AX601810 Sequence
39	155.6	84.1	2084	6	BD077082 Alpha-amy
40	155.6	84.1	2084	6	BD085805 Alpha-amy
41	155.6	84.1	2604	6	AR168304 Sequence
42	155.6	84.1	2604	6	BD249244 Alpha-amy
43	155.6	84.1	2604	6	AR215288 Sequence
44	155.6	84.1	2604	6	AX036892 Sequence
45	155.6	84.1	2604	6	BD137366 Enzymatic

ALIGNMENTS

RESULT 1
AR160502
LOCUS AR160502 Sequence 26 from patent US 6255076. 185 bp DNA linear PAT 17-OCT-2001
DEFINITION
ACCESSION AR160502
VERSION AR160502.1 GI:16224477
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner,W., Sloma,A. and Thomas,M.D.
TITLE Methods for producing a polypeptide in a Bacillus cell
JOURNAL Patent: US 6255076-A 26 03-JUL-2001;
FEATURES Location/Qualifiers

```
source 1. .185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 7.2e-35;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGAGGTTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTAATTTTACTGATATGTAATAATTTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTGACATTAATTTTACTGATATGTAATAATTTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 2
AX088920
LOCUS AX088920 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 21 from Patent WO0114534.
ACCESSION AX088920
VERSION AX088920.1 GI:13397679
SOURCE Bacillus subtilis
ORGANISM Bacillus subtilis
REFERENCE 1
AUTHORS Thomas, M.D. and Brown, K.M.
TITLE Polypeptides having pectin acetyltransferase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 21 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source 1. .185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 7.2e-35;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGAGGTTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTAATTTTACTGATATGTAATAATTTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTGACATTAATTTTACTGATATGTAATAATTTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
BD136798
LOCUS BD136798 185 bp DNA linear PAT 18-SEP-2002
```

```
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136798
VERSION BD136798.1 GI:23231743
KEYWORDS JP 2002504379-A/26.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner, W., Sloma, A. and Thomas, M.D.
TITLE Process for producing polypeptide in bacillus cells
JOURNAL Patent: JP 2002504379-A 26 12-FEB-2002;
NOVO NORDISK BIOTECH INC
COMMENT OS Bacillus
PN JP 2002504379-A/26
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/09, C12N1/21, C12N9/54, C07K14/325, C12N1/21, C12R1/07,
PC C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source 1. .185
/organism="Bacillus"
/organism="Bacillus"
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 7.2e-35;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACACCAATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGAGGTTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTAATTTTACTGATATGTAATAATTTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTGACATTAATTTTACTGATATGTAATAATTTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
AR160503
LOCUS AR160503 185 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 27 from patent US 6255076.
ACCESSION AR160503
VERSION AR160503.1 GI:16224480
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner, W., Sloma, A. and Thomas, M.D.
TITLE Methods for producing a polypeptide in a Bacillus cell
JOURNAL Patent: US 6255076-A 27 03-JUL-2001;
FEATURES
source 1. .185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
```

```
Best Local Similarity 99.5%; Pred. No. 1.8e-34; Mismatches 1; Indels 0; Gaps 0;
Matches 184; Conservative 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAAGGA 120
Db 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTTTACTGATGTATATAATATTTGTATAAGAAATGG 180
Db 121 ATAAAGGGGGTTCACATTTTACTGATGTATATAATATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 5
AX088921
LOCUS AX088921 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 22 from Patent WO0114534.
ACCESSION AX088921
VERSION AX088921.1 GI:13397680
KEYWORDS Bacillus subtilis
SOURCE Bacillus subtilis
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Thomas, M.D. and Brown, K.M.
TITLE Polypeptides having pectin acetyltransferase activity and nucleic
JOURNAL acids encoding same
Patent: WO 0114534-A 22 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source
1. 185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"
ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
Best Local Similarity 99.5%; Pred. No. 1.8e-34;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAAGGA 120
Db 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTTTACTGATGTATATAATATTTGTATAAGAAATGG 180
Db 121 ATAAAGGGGGTTCACATTTTACTGATGTATATAATATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 6
BD136799
LOCUS BD136799 185 bp DNA linear PAT 18-SEP-2002
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136799
VERSION BD136799.1 GI:23231744
KEYWORDS JP 2002504379-A/27.
SOURCE unidentified
ORGANISM unidentified
```

```
unclassified.
1 (bases 1 to 185)
Widner, W., Sloma, A. and Thomas, M.D.
Process for producing polypeptide in bacillus cells
Patent: JP 2002504379-A 27 12-FEB-2002;
NOVO NORDISK BIOTECH INC
OS Bacillus
PN JP 2002504379-A/27
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/00, C12N1/21, C12N9/54//C07K14/325, (C12N1/21, C12R1:07),
PC C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source
1. 185
Location/Qualifiers
/organism="Bacillus".
FEATURES
source
1. 185
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
Best Local Similarity 99.5%; Pred. No. 1.8e-34;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAAGGA 120
Db 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTTTACTGATGTATATAATATTTGTATAAGAAATGG 180
Db 121 ATAAAGGGGGTTCACATTTTACTGATGTATATAATATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 7
AR160501
LOCUS AR160501 185 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 25 from patent US 6255076.
ACCESSION AR160501
VERSION AR160501.1 GI:16224474
KEYWORDS SOURCE Unknown.
ORGANISM Unknown.
unclassified.
1 (bases 1 to 185)
Widner, W., Sloma, A. and Thomas, M.D.
Methods for producing a polypeptide in a Bacillus cell
Patent: US 6255076-A 25 03-JUL-2001;
LOCATION/Qualifiers
1. 185
/organism="unknown"
/mol_type="unassigned DNA"
source
ORIGIN
Query Match 96.5%; Score 178.6; DB 6; Length 185;
Best Local Similarity 97.8%; Pred. No. 2.6e-33;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
```

```
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTATTTACTGATATGTAATAATTTGTAAGAATGG 180
Db 121 ATAAAGGGGGTGTATTATTTACTGATATGTAATAATTTGTAAGAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 8
AX088919
LOCUS 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 20 from Patent WO0114534.
ACCESSION AX088919
VERSION AX088919.1 GI:13397678
KEYWORDS Bacillus subtilis
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Thomas, M.D. and Brown, K.M.
TITLE Polypeptides having pectin acetyltransferase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 20 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
Source
1. .185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 96.5%; Score 178.6; DB 6; Length 185;
Best Local Similarity 97.8%; Pred. No. 2.6e-33;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGACCAATATAAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGACCAATATAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTATTTACTGATATGTAATAATTTGTAAGAATGG 180
Db 121 ATAAAGGGGGTGTATTATTTACTGATATGTAATAATTTGTAAGAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 9
BD136797
LOCUS 185 bp DNA linear PAT 18-SEP-2002
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136797
VERSION BD136797.1 GI:23231742
KEYWORDS unidentified
SOURCE unclassified
ORGANISM 1 (bases 1 to 185)
REFERENCE 1
AUTHORS Widner, W., Sloma, A. and Thomas, M.D.
TITLE Process for producing polypeptide in bacillus cells
JOURNAL Patent: JP 2002504379-A 25 12-FEB-2002;
NOVO NORDISK BIOTECH INC

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTATTTACTGATATGTAATAATTTGTAAGAATGG 180
Db 121 ATAAAGGGGGTGTATTATTTACTGATATGTAATAATTTGTAAGAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 10
I40596
LOCUS 162 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 24 from patent US 5621089.
ACCESSION I40596
VERSION I40596.1 GI:2082888
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 162)
AUTHORS Sloma, A.P., Outtrup, H., Dammann, C. and Aaslyng, D.A.
TITLE Nucleic acid constructs for the production of a Bacillus alkaline protease
JOURNAL Patent: US 5621089-A 24 15-APR-1997;
FEATURES
Source
1. .162
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 84.1%; Score 155.6; DB 6; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.1e-27;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCCGATTGTTTGAGAAAAGAGAGACCAATATAAATACCTTGTGTATCAGACAGGT 77
Db 1 ATCCGATTGTTTGAGAAAAGAGAGACCAATATAAATACCTTGTGTATCAGACAGGT 60
QY 78 ATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGAATATAAGGGGGTTCACA 137
Db 61 ATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGAATATAAGGGGGTTCACA 120
```

```
COMMENT OS Bacillus
PN JP 2002504379-A/25
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/09, C12N1/21, C12N9/54//C07K14/325, (C12N1/21, C12R1:07),
CC C12N15/00
Process for producing polypeptide in bacillus cells FH Key
Location/Qualifiers
source 1. .185
/organism="Bacillus".
FEATURES
Source
1. .185
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 96.5%; Score 178.6; DB 6; Length 185;
Best Local Similarity 97.8%; Pred. No. 2.6e-33;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGACCAATATAAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGACCAATATAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTATTTACTGATATGTAATAATTTGTAAGAATGG 180
Db 121 ATAAAGGGGGTGTATTATTTACTGATATGTAATAATTTGTAAGAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 10
I40596
LOCUS 162 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 24 from patent US 5621089.
ACCESSION I40596
VERSION I40596.1 GI:2082888
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 162)
AUTHORS Sloma, A.P., Outtrup, H., Dammann, C. and Aaslyng, D.A.
TITLE Nucleic acid constructs for the production of a Bacillus alkaline protease
JOURNAL Patent: US 5621089-A 24 15-APR-1997;
FEATURES
Source
1. .162
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 84.1%; Score 155.6; DB 6; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.1e-27;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCCGATTGTTTGAGAAAAGAGAGACCAATATAAATACCTTGTGTATCAGACAGGT 77
Db 1 ATCCGATTGTTTGAGAAAAGAGAGACCAATATAAATACCTTGTGTATCAGACAGGT 60
QY 78 ATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGAATATAAGGGGGTTCACA 137
Db 61 ATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGAATATAAGGGGGTTCACA 120
```

```

Qy 138 TTATTTTACTGATATGATATATATATATATATTTCTGTATTAAGAAAATG 179
Db 121 TTATTTTACTGATATGTAATAATATATATTTCTGTATTAAGAAAATG 162

RESULT 11
LOCUS I40867 162 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 24 from patent US 5622841.
ACCESSION I40867
VERSION I40867.1 GI:2082347
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 162)
AUTHORS Sloma,A.P., Outtrup,H., Dambmann,C. and Aaslyng,D.A.
TITLE Method for the production of heterologous polypeptides using a
promoter element and signal peptide of a bacillus gene encoding an
alkaline protease
JOURNAL Patent: US 5622841-A 24 22-APR-1997;
FEATURES Location/Qualifiers
source 1..162
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 84.1%; Score 155.6; DB 6; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.1e-27;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGG 77
Db 1 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGG 60

Qy 78 ATTTTATGCTGCCAGACTGCCGCTGTGTAAAAAAGGATAAAGGGGGTTGACA 137
Db 61 ATTTTATGCTGCCAGACTGCCGCTGTGTAAAAAAGGATAAAGGGGGTTGTTA 120

Qy 138 TTATTTTACTGATATGATATATATATATTTCTGTATTAAGAAAATG 179
Db 121 TTATTTTACTGATATGTAATAATATATATTTCTGTATTAAGAAAATG 162

RESULT 12
LOCUS I40895 162 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 24 from patent US 5622850.
ACCESSION I40895
VERSION I40895.1 GI:2082375
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 162)
AUTHORS Sloma,A.P., Outtrup,H., Dambmann,C. and Aaslyng,D.A.
TITLE Recombinant methods for the production of a bacillus alkaline
protease
JOURNAL Patent: US 5622850-A 24 22-APR-1997;
FEATURES Location/Qualifiers
source 1..162
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 84.1%; Score 155.6; DB 6; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.1e-27;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGG 77
Db 1 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGG 60

Qy 78 ATTTTATGCTGCCAGACTGCCGCTGTGTAAAAAAGGATAAAGGGGGTTGACA 137
Db 61 ATTTTATGCTGCCAGACTGCCGCTGTGTAAAAAAGGATAAAGGGGGTTGTTA 120

Qy 138 TTATTTTACTGATATGATATATATATATTTCTGTATTAAGAAAATG 179
Db 121 TTATTTTACTGATATGTAATAATATATATTTCTGTATTAAGAAAATG 162

```

```

Qy 78 ATTTTATGCTGCCAGACTGCCGCTGTGTAAAAAAGGATAAAGGGGGTTGACA 137
Db 61 ATTTTATGCTGCCAGACTGCCGCTGTGTAAAAAAGGATAAAGGGGGTTGTTA 120

Qy 138 TTATTTTACTGATATGATATATATATTTCTGTATTAAGAAAATG 179
Db 121 TTATTTTACTGATATGTAATAATATATTTCTGTATTAAGAAAATG 162

RESULT 13
LOCUS I56837 162 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 24 from patent US 5650326.
ACCESSION I56837
VERSION I56837.1 GI:2477250
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 162)
AUTHORS Sloma,A.P., Outtrup,H., Dambmann,C. and Aaslyng,D.Anita.
TITLE Promoter element and signal peptide of a gene encoding a Bacillus
alkaline protease and vectors comprising same
JOURNAL Patent: US 5650326-A 24 22-JUL-1997;
FEATURES Location/Qualifiers
source 1..162
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 84.1%; Score 155.6; DB 6; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.1e-27;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGG 77
Db 1 ATCGATTGTTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCATCAGACAGG 60

Qy 78 ATTTTATGCTGCCAGACTGCCGCTGTGTAAAAAAGGATAAAGGGGGTTGACA 137
Db 61 ATTTTATGCTGCCAGACTGCCGCTGTGTAAAAAAGGATAAAGGGGGTTGTTA 120

Qy 138 TTATTTTACTGATATGATATATATATTTCTGTATTAAGAAAATG 179
Db 121 TTATTTTACTGATATGTAATAATATATTTCTGTATTAAGAAAATG 162

RESULT 14
LOCUS BD174625 249 bp DNA linear PAT 18-MAR-2003
DEFINITION Modified promoter.
ACCESSION BD174625
VERSION BD174625.1 GI:29120315
KEYWORDS JP 2002272466-A/14.
SOURCE Bacillus amyloliquefaciens
ORGANISM Bacillus amyloliquefaciens
REFERENCE 1 (bases 1 to 249)
AUTHORS Mizubuchi,H., Fushimi,N. and Miyoshi,S.
TITLE Modified promoter
JOURNAL Patent: JP 2002272466-A 14 24-SEP-2002;
COMMENT SHOWA SANGYO CO LTD
OS Bacillus amyloliquefaciens
PN JP 2002272466-A/14
PD 24-SEP-2002
PF 18-MAR-2001 JP 2001074780
PI HROVUKI MIZUBUCHI,NAOYA FUSHIMI,SHINSUKE MIYOSHI PC
C12N15/09,C12N1/21,C12P21/02//C12N1/21,C12R1/07),C12N1/21,PC
C12R1:01),
PC (C12N1/21,C12R1:19),C12N15/00
CC Modified promoter
EH Key Location/Qualifiers
FT source 1..249

```

```

FT      /organism='Bacillus amyloliquefaciens'.
FEATURES             Location/Qualifiers
     source            1..249
                        /organism="Bacillus amyloliquefaciens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:1390"

ORIGIN
Query Match      84.1%; Score 155.6; DB 6; Length 249;
Best Local Similarity 97.5%; Pred. No. 9.7e-28;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTCTGTCATCAGACAGGTT 77
    |||||
DB 73 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTCTGTCATCAGACAGGTT 132
    |||||

QY 78 ATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAAGGAATAAAGGGGGTTGACA 137
    |||||
DB 133 ATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAAGGAATAAAGGGGGTTGTTA 192
    |||||

QY 138 TTATTTTACTGATGTATATAATATTTGTATAGAAAATG 179
    |||||
DB 193 TTATTTTACTGATGTATATAATATTTGTATAGAAAATG 234
    |||||

RESULT 15
BAMYL
LOCUS      537 bp DNA linear BCT 04-APR-1995
DEFINITION Bacillus amyloliquefaciens gene fragment encoding alpha-amylase.
ACCESSION V00092
VERSION    V00092.1 GI:39297
KEYWORDS   amylase; alpha; signal peptide.
SOURCE     Bacillus amyloliquefaciens
            Bacillus amyloliquefaciens
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 537)
AUTHORS    Palva, I., Pettersson, R.F., Kalkkinen, N., Lehtovaara, P., Sarvas, M.,
            Soderlund, H., Takkinen, K. and Kaariainen, L.
TITLE      Nucleotide sequence of the promoter and NH2-terminal signal peptide
            region of the alpha-amylase gene from Bacillus amyloliquefaciens
JOURNAL    Gene 15 (1), 43-51 (1981)
MEDLINE    82051296
PUBMED     6170539
FEATURES             Location/Qualifiers
     source            1..537
                        /organism="Bacillus amyloliquefaciens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:1390"
                        250..>537
                        /note="unnamed protein product; amylase"
                        /codon_start=1
                        /transl_table=11
                        /protein_id="CAA23430.1"
                        /db_xref="GI:39298"
                        /db_xref="GOA:P00692"
                        /db_xref="SWISS-PROT:P00692"
                        /translation="MIQKARTVSRFLVLMCTLLFVSLPTTKTSVAVNGILMQYFHWY
                        PNDQHWKRLQNDAEHLSDIGITAVWIPPAKGLSOSDNGYGPYDLDLGEF"
                        250..342
                        sig_peptide
                        /note="signal peptide"

ORIGIN
Query Match      84.1%; Score 155.6; DB 1; Length 537;
Best Local Similarity 97.5%; Pred. No. 8.3e-28;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTCTGTCATCAGACAGGTT 77
    |||||
DB 73 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTCTGTCATCAGACAGGTT 132
    |||||

QY 78 ATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAAGGAATAAAGGGGGTTGACA 137
    |||||
```

```

Db 133 ATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAAGGAATAAAGGGGGTTGTTA 192
QY 138 TTATTTTACTGATGTATATAATATTTGTATAGAAAATG 179
    |||||
Db 193 TTATTTTACTGATGTATATAATATTTGTATAGAAAATG 234
    |||||
```

Search completed: March 19, 2004, 02:14:35
Job time : 1081.5 secs